

# The Journal of PSYCHOLOGY

The General Field of Psychology

POWELL MURCHISON  
*Managing Editor*

HELEN M. DAVIS  
*Associate Editor*

## *Editorial Board*

JOHN A. ANTOINETTI  
Life Insurance Agency  
Management Association

ARTHUR G. BILLS  
University of Cincinnati

E. A. BOTT  
University of Toronto

LEONARD CARMICHAEL  
National Geographic Society

HERBERT S. CONRAD  
U. S. Office of Education

JOHN F. DASHIELL  
Florida State University

LEONARD W. DOOB  
Yale University

ROY M. DORCUS  
University of California  
at Los Angeles

FRANK A. GELDARD  
Princeton University

If this space should be unstamped, this is the regular library edition.  
But if this space is stamped with a designating title, this is a special

edition, sold under the restrictions of a bilateral contract, and may not  
be resold for a period of five years from date of publication.

CLARENCE H. GRAHAM  
Columbia University

ERNEST R. HILGARD  
Stanford University

JOHN E. HORROCKS  
Ohio State University

HEINRICH KLÜVER  
University of Chicago

NORMAN L. MUNN  
University of Adelaide,  
Australia

WARREN H. TEICHNER  
Northeastern University

LEE EDWARD TRAVIS  
University of Southern  
California

ROBERT C. TRYON  
University of California

JOSEPH ZUBIN  
Psychiatric Institute,  
Columbia University

## VOLUME 71

### 1969

Copyright © 1969 by The Journal Press  
Provincetown, Massachusetts, U. S. A.

All rights reserved

Published at 35 New St., Worcester, Mass.

and all mail to Managing Editor, 2 Commercial Street, Provincetown, Massachusetts 02657

# AUTHOR INDEX

Al-Issa, Ihsan .....	143	Kline, Milton V. ....	21
Angelino, Henry .....	253	LaFaro, Dolores .....	155
Attkisson, Clifford .....	27	Lindauer, M. S. ....	41
Ault, Ruth L. ....	191	Linder, Morton .....	21
Ball, Richard J. ....	185	Merlis, Sidney .....	127
Bartley, S. Howard .....	185, 281	Mills, John A. ....	69
Bentler, P. M. ....	33	Monty, Richard A. ....	111
Berger, Vincent F. ....	253	Moroney, William F. ....	77
Bottrill, John .....	89	Moseley, Nicholas .....	3
Braun, John R. ....	155	Munz, David C. ....	253
Cassel, Russell N. ....	199	O'Neill, Brian J. ....	95
Collins, Lois .....	127	Pedersen, Darhl M. ....	225
Comrey, Andrew L. ....	45	Phelan, Joseph G. ....	205
Cutts, Norma E. ....	3	Richardson, Ed. ....	205
DeFazio, Victor J. ....	77	Schludermann, Eduard .....	259
Dreistadt, Roy .....	159	Schludermann, Shirin .....	259
Eichsteadt, Arden C. ....	199	Sciortino, Rio .....	261, 271
Engel, Gerald .....	133	Sheppard, Charles .....	127
Fiorentino, Diane .....	127	Shrader, Raymond R. ....	27
Fisher, Dennis F. ....	111	Smouse, Albert D. ....	253
Gilley, Hoyt Melvyn .....	245	Stanley, Gordon .....	83
Glucksberg, Sam .....	111	Vantress, Florence E. ....	59
Goss, Allen .....	17	Vitulli, William F. ....	177
Gunderson, Maxine Murphy .....	217	Vogler, Roger E. ....	191
Handler, Leonard .....	27	Wessler, Richard L. ....	63
Jackson, Robert .....	83	Williams, Christine Blanton .....	59
Jamison, Kay .....	45	Winocur, Gordon .....	69
Kazsuk, Jonathan D. ....	281	Yarmey, A. Daniel .....	95

# TABLE OF CONTENTS

Notes on photographic memory . . . . .	3
BY NORMA E. CUTTS AND NICHOLAS MOSELEY	
Abasement scores and adjustment of neuropsychiatric patients . . . . .	17
BY ALLEN GOSS	
Psychodynamic factors in the experimental investigation of hypnotically induced emotions with particular reference to blood glucose measurements . . . . .	21
BY MILTON V. KLINE AND MORTON LINDER	
The use of figure drawings to assess religious values . . . . .	27
BY CLIFFORD ATTKISSON, LEONARD HANDLER, AND RAYMOND R. SHRADER	
Semantic space is (approximately) bipolar . . . . .	33
BY P. M. BENTLER	
Historical and contemporary attitudes toward numbers . . . . .	41
BY M. S. LINDAUER	
A comparison of personality factor structure in British and American university students . . . . .	45
BY KAY JAMISON AND ANDREW L. COMREY	
Relation between internal-external control and aggression . . . . .	59
BY CHRISTENE BLANTON WILLIAMS AND FLORENCE E. VANTRESS	
Experimenter expectancy effects in three dissimilar tasks . . . . .	63
BY RICHARD L. WESSLER	
Retroactive inhibition of connected discourse as a function of meaningfulness and of degree of original learning . . . . .	69
BY JOHN A. MILLS AND GORDON WINOCUR	
Performance characteristics of field dependent and independent individuals on an auditory signal detection task . . . . .	77
BY VICTOR J. DEFazio AND WILLIAM F. MORONEY	
The effect of leading contour on the relative lengths of moving light and dark arcs . . . . .	83
BY GORDON STANLEY AND ROBERT JACKSON	
An investigation of some intrinsic variables affecting test responses . . . . .	89
BY JOHN BOTTRILL	
S-R and R-S paired-associate learning as a function of concreteness, imagery, specificity, and association value . . . . .	95
BY A. DANIEL YARMEY AND BRIAN J. O'NEILL	
Visual confusion matrices: Fact or artifact? . . . . .	111
BY DENNIS F. FISHER, RICHARD A. MONTY, AND SAM GLUCKSBERG	
Further study of performance errors on Ravens Progressive Matrices (1938) . . . . .	127
BY CHARLES SHEPPARD, DIANE FIORENTINO, LOIS COLLINS, AND SIDNEY MERLIS	
Comparison between American permanent residents of Israel: Part I, American background . . . . .	133
BY GERALD ENGEL	



Problems in the cross-cultural study of schizophrenia . . . . .	143
By IHSAN AL-ISBA	
Faking and faking detection on the 16 PF-Form A . . . . .	155
By JOHN R. BRAUN AND DOLORES LAFARO	
The use of analogies and incubation in obtaining insights in creative problem solving . . . . .	159
By ROY DREISTADT	
The effects of punishment on food pellet discrimination . . . . .	177
By WILLIAM F. VITULLI	
Visual acuity decrement for a subfusional intermittency with two directions of border contrast . . . . .	185
By S. HOWARD BARTLEY AND RICHARD J. BALL	
Problem-solving motivation in verbal conditioning studies . . . . .	191
By ROGER E. VOGLER AND RUTH L. AULT	
Factorial structure of CQT, ACT, and SAT test scores for 50 available college freshmen . . . . .	199
By RUSSELL N. CASSEL AND ARDEN C. EICHSTEADT	
Cognitive complexity, strategy of the other player, and two-person game behavior . . . . .	205
By JOSEPH G. PHELAN AND ED. RICHARDSON	
Relationships between expressed personality needs and social background and military status variables . . . . .	217
By MAXINE MURPHY GUNDERSON	
Evaluation of self and others and some personality correlates . . . . .	225
By DARHL M. PEDERSEN	
Effects of vicarious verbal stimuli on conditioning of hostile and neutral verbs . . . . .	245
By HOYT MELVYN GILLEY	
The effects of item difficulty sequencing and anxiety reaction type on aptitude test performance . . . . .	253
By VINCENT F. BERGER, DAVID C. MUNZ, ALBERT D. SMOUSE, AND HENRY ANGELINO	
A methodological note on "a self-concept inventory" . . . . .	259
By SHIRIN SCHLUDERMANN AND EDUARD SCHLUDERMANN	
Factorial study and analysis of factor variance of intellective self-ratings from a combined sample of male and female subjects . . . . .	261
By RIO SCIORTINO	
Factorial study of general adaptability self-ratings by male and female subjects . . . . .	271
By RIO SCIORTINO	
Brightness matching under intermittent stimulation . . . . .	281
By JONATHAN D. KAZSUK AND S. HOWARD BARTLEY	



# Journal of Genetic Psychology

Founded by G. Stanley Hall in 1891; reorganized by Carl Murchison in 1925

Quarterly, published in March, June, September, and December. Devoted to child behavior, animal behavior, and the problems of aging. Manuscripts may be sent to any member of the Editorial Board, or may be sent directly to The Journal Press. All subscriptions and business communications should be sent directly to The Journal Press. Two volumes annually. Per annum, \$20.00; per volume, \$15.00; single numbers, \$7.50. Complete sets from 1891 at \$15.00 per volume, plus transportation.

---

# Journal of General Psychology

Founded in 1927 by E. B. Titchener and Carl Murchison

Quarterly, published in January, April, July, and October. Devoted primarily to experimental, theoretical, clinical, and historical psychology. Manuscripts may be sent to any member of the Editorial Board, or may be sent directly to The Journal Press. All subscriptions and business communications should be sent directly to The Journal Press. Two volumes annually. Per annum, \$20.00; per volume, \$15.00; single numbers, \$7.50. Complete sets from 1927 at \$15.00 per volume, plus transportation.

---

# Journal of Social Psychology

Founded in 1929 by John Dewey and Carl Murchison

Bimonthly, beginning with February. National, group, cultural, racial, and differential psychology. Manuscripts may be sent to any member of the Editorial Board, or may be sent directly to The Journal Press. All subscriptions and business communications should be sent directly to The Journal Press. Bimonthly, three volumes per year; per annum, \$30.00; per volume, \$15.00; single numbers, \$7.50. First number is dated February, 1930. Complete sets from 1930 at \$15.00 per volume, plus transportation.

---

*Each author will receive 100 unbound reprints gratis. Additional reprints may be purchased at a nominal cost when the author returns page-proof.*

THE JOURNAL PRESS  
2 Commercial Street  
Provincetown, Massachusetts  
U. S. A.

# Club Rate for Five International Psychological Journals

	PER ANNUM
THE JOURNAL OF GENETIC PSYCHOLOGY . . . . . (Quarterly, two volumes per year)	\$20.00
GENETIC PSYCHOLOGY MONOGRAPHS . . . . . (Quarterly, two volumes per year)	\$20.00
THE JOURNAL OF GENERAL PSYCHOLOGY . . . . . (Quarterly, two volumes per year)	\$20.00
THE JOURNAL OF SOCIAL PSYCHOLOGY . . . . . (Bimonthly, three volumes per year)	\$30.00
THE JOURNAL OF PSYCHOLOGY . . . . . (Bimonthly, three volumes per year)	\$30.00
	<hr/>
TOTAL	\$120.00
TEACHERS AND STUDENTS EDITION, <i>Restricted Sale</i>	\$50.00

The *Teachers and Students Edition* is open to any individual psychologist, teacher, student, or physician but requires an agreement that the edition will not be offered for sale till after five years from publication and an agreement that the edition will not be substituted for an institutional subscription.

## THE JOURNAL PRESS

2 Commercial Street  
Provincetown, Massachusetts  
U. S. A.

## Genetic Psychology Monographs

Founded in 1925 by Carl Murchison

Quarterly: February, May, August, November. Two volumes per year. Child behavior, animal behavior, and the problems of aging. Each number one or more complete researches. Manuscripts may be sent to any member of the Editorial Board, or may be sent directly to The Journal Press. All subscriptions and business communications should be sent directly to The Journal Press. Two volumes annually. Per annum, \$20.00; per volume, \$15.00; single numbers, \$7.50. Complete sets from 1926 at \$15.00 per volume, plus transportation.

**NOTICE.** Effective immediately we offer prior publication for certain selected monographs. Charges will be based on the announced rates for *The Journal of Psychology*. For a limited period of time our policy is changed for regular monographs: Authors will be charged for cuts and expensive tables, but no page charge will be imposed. Two hundred reprints of each monograph will be supplied gratis.

The listing of monographs for 1926 through 1963 will be printed in the first half of each volume of all five journals. Monographs published after 1963 will be listed in the second half of each volume.



# Genetic Psychology Monographs (continued)

## VOLUME 1-1926

1. Performance tests for children of pre-school age—R. STUTSMAN
2. An experimental study of the eidetic type—H. KLÜVER
- 3 & 4. A study of natio-racial mental differences—N. D. M. HIRSCH
5. A psychological study of juvenile delinquency by group methods—J. W. BRIDGES AND K. M. B. BRIDGES
6. The influence of puberty praecox upon mental growth—A. GSELL

## VOLUME 2-1927

- 1 & 2. The mind of a gorilla—R. M. YERKES
3. The role of eye-muscles and mouth-muscles in the expression of the emotions—K. DUNLAP
4. Family similarities in mental-test abilities—R. R. WILLOUGHBY
5. Coordination in the locomotion of infants—L. H. BURNISSE
6. The mind of a gorilla: Part II. Mental development—R. M. YERKES

## VOLUME 3-January-June, 1928

1. An experimental study of the olfactory sensitivity of the white rat—J. R. LIGGETT
2. A photographic study of eye movements in reading formulae—M. A. TINKER
3. An experimental study of the East Kentucky mountaineers—N. D. M. HIRSCH
4. Responses of foetal guinea pigs prematurely delivered—G. T. AVERY
5. Objective differentiation between three groups in education (teachers, research workers, and administrators)—M. B. JENKINS
6. The effect of segregation on the sex behavior of the white rat as measured by the obstruction method—M. JENKINS

## VOLUME 4-July-December, 1928

1. Observation and training of fundamental habits in young children—E. A. BOTT, W. E. BLATZ, N. CHANT, AND H. BOTT
- 2 & 3. Determination of a content of the course in literature of a suitable difficulty for junior and senior high school students—M. C. BURCH
- 4 & 5. Methods for diagnosis and treatment of cases of reading disability—M. MONROE
6. The relative effectiveness of lecture and individual reading as methods of college teaching—E. B. GREENE

## VOLUME 5-January-June, 1929

1. The age factor in animal learning: I. Rats in the problem box and the maze—C. P. STONE
2. The effect of delayed incentive on the hunger drive in the white rat—E. L. HAMILTON
3. Which hand is the eye of the blind?—J. M. SMITH
4. The effect of attitude on free word association-time—A. G. EKDAHL
5. The localization of tactual space: A study of average and constant errors under different types of localization—L. E. COLE
6. The effects of gonadectomy, vasotomy, and injections of placental and orchic extracts on the sex behavior of the white rat—H. W. NISSEN

## VOLUME 6-July-December, 1929

1. Learning and growth in identical infant twins: An experimental study by the method of co-twin control—A. GSELL AND H. THOMPSON
2. The age factor in animal learning: II. Rats on a multiple light discrimination box and a difficult maze—C. P. STONE
3. The acquisition and interference of motor habits in young children—E. MCGINNIS
4. A vocational and socio-educational survey of graduates and non-graduates of small high schools of New England—A. D. MUELLER
- 5 & 6. A study of the smiling and laughing of infants in the first year of life—R. W. WASHBURN

## VOLUME 7-January-June, 1930

1. Tensions and emotional factors in reaction—E. DUFFY
2. Teacher influence on class achievement: A study of the relationship of estimated teaching ability to pupil achievement in reading and arithmetic—H. R. TAYLOR
- 3 & 4. A study of the effect of inverted retinal stimulation upon spatially coordinated behavior—P. H. EWERT
5. A study of the effect of inverted retinal stimulation upon spatially coordinated behavior—E. F. LOND
6. An experimental study upon three hundred school children over a six-year period—N. D. M. HIRSCH

## VOLUME 8-July-December, 1930

1. The amount and nature of activities of newborn infants under constant external stimulating conditions during the first ten days of life—O. C. IRWIN
2. Race and social differences in performance tests—S. D. PORTER, *et al.*
3. Language and growth: The relative efficacy of early and deferred vocabulary training, studied by the method of co-twin control—L. C. STRAYER
4. Eye-movements and optic nystagmus in early infancy—J. M. MCGINNIS
- 5 & 6. Reactions of kindergarten, first-, and second-grade children to constructive play materials—L. FARWELL

## VOLUME 9-January-June, 1931

- 1 & 2. The status of the first-born with special reference to intelligence—H. H. HSIAO
- 3 & 4. An experimental study of bright, average, and dull children at the four-year mental level—H. P. DAVIDSON
5. An historical, critical, and experimental study of the Seashore-Kwalwasser test battery—P. R. FARNSWORTH
6. A comparison of difficulty and improvement in the learning of bright and dull children in reproducing a descriptive selection—F. T. WILSON

## VOLUME 10-July-December, 1931

1. A comparative study of a group of southern white and negro infants—M. B. MCGRAW
- 2 & 3. An experimental study of prehension in infants by means of systematic cinema records—H. M. HALVERSON
4. The limits of learning ability in kittens—A. M. SHUEV
- 5 & 6. The effect of habit interference upon performance in maze learning—O. W. ALM

## VOLUME 11-January-June, 1932

1. General factors in transfer of training in the white rat—T. A. JACKSON
2. The effect of color on visual apprehension and perception—M. A. TINKER
3. The reliability and validity of maze experiments with white rats—R. LEPPER
4. A critical study of two lists of best books for children—F. K. SHUTTLEWORTH
- 5 & 6. Measuring human energy cost in industry: A general guide to the literature—R. M. PAGE

## VOLUME 12-July-December, 1932

1. Family resemblances in verbal and numerical abilities—H. D. CARTER
2. The development of fine prehension in infancy—B. M. CASTNER
- 3 & 4. The growth of adaptive behavior in infants: An experimental study at seven age levels—H. M. RICHARDSON
- 5 & 6. Differential reactions to taste and temperature stimuli in newborn infants—K. JENSEN

## VOLUME 13-January-June, 1933

1. A critique of sublimation in males: A study of forty superior single men—H. L. KOCH, *et al.*
2. A study of the nature, measurement, and determination of hand preference—H. L. KOCH, *et al.*
3. The growth and decline of intelligence: A study of a homogeneous group between the ages of ten and sixty—H. E. JONES AND H. S. CONRAD
4. The relation between the complexity of the habit to be acquired and the form of the learning curve in young children—M. L. MATTHEWSON
5. Eating habits in relation to personality development of two- and three-year-old children: A study of sixty-nine children in two nursery schools—A. A. ELLIOT
6. Coordinating mechanisms of the spinal cord—O. C. INGEBRETSEN



# Genetic Psychology Monographs (continued)

VOLUME 14—July-December, 1933

1. Mental growth during the first three years: A developmental study of sixty-one children by repeated tests—N. BAYLEY
2. A study of triplets: including theories of their possible genetic relationships—F. N. ANDERSON AND N. V. SCHEIDEMANN
3. The objective measurement of emotional reactions—H. V. GASKILL
4. Development of behavior in the fetal cat—J. D. CORONIOS
5. A study of certain language developments of children in grades four to twelve, inclusive—L. L. LABRANT
6. The effect of early and delayed practice on memory and motor performances studied by the method of co-twin control—J. R. HILGARD

VOLUME 15—January-June, 1934

1. Studies in the psychology of tone and music—P. R. FARNSWORTH
2. Motor learning of children in equilibrium in relation to nutrition—E. L. BEEBE
3. Discrimination limits of pattern and size in the goldfish *Carassius auratus*—J. B. ROWLEY
4. Limits of learning ability in the white rat and the guinea pig—B. F. RIES
- 5 & 6. The limits of learning ability in rhesus monkeys—H. A. FJELD

VOLUME 16—July-December, 1934

1. A statistical study of ratings on the California Behavior Inventory for Nursery-School Children—H. S. CONRAD
2. An eye-movement study of objective examination questions—A. FRANDSEN
3. An experimental study of constitutional types—O. KLINEBERG, S. E. ASCH, AND H. BLOCK
4. The development of a battery of objective group tests of manual laterality, with the results of their application to 1300 children—W. N. DUROST
- 5 & 6. An experimental study in the prenatal guinea-pig of the origin and development of reflexes and patterns of behavior in relation to the stimulation of specific receptor areas during the period of active fetal life—L. CARMICHAEL

VOLUME 17—January-December, 1935

1. Organization of behavior in the albino rat—R. L. THORNDIKE
2. Brightness discrimination in the rhesus monkey—M. P. CRAWFORD
3. The limits of learning ability in cebus monkeys—A. M. KOCH
4. Nature-nurture and intelligence—A. M. LEAHY
5. On intelligence of epileptic children—E. B. SULLIVAN AND L. GAHAGAN
6. A study of the play of children of preschool age by an unobserved observer—D. L. COCKRELL

VOLUME 18—January-December, 1936

1. Sex differences in variational tendency—Q. MCNEMAR AND L. M. TERMAN
2. The process of learning to dress among nursery-school children—C. B. KEY, M. R. WHITE, M. P. HONZIK, A. B. HEINEY, AND D. ERWIN
3. A study of the present social status of a group of adults, who, when they were in elementary schools, were classified as mentally deficient—W. R. BAKER
4. The influence of specific experience upon mental organizations—A. ANASTASI
- 5 & 6. Studies in aggressiveness—L. BENDER, S. KEISER, AND P. SCHILDER

VOLUME 19—January-December, 1937

1. Psychological bases of self-mutilation—C. DABROWSKI
2. Masculine temperament and secondary sex characteristics: A study of the relationship between psychological and physical measures of masculinity—H. GILKINSON
3. A psychological study of forty unmarried mothers—R. D. NOTTINGHAM
4. Behavior problems in the children of psychotic and criminal parents—L. BENDER
5. Domination and integration in the social behavior of young children in an experimental play situation—H. H. ANDERSON
6. The sequential patterning of prone progression in the human infant—L. B. AMES

VOLUME 20—January-December, 1938

1. The relationship between characteristics of personality and physique in adolescents—P. S. DE Q. CAROT
2. Behavior problems of elementary school children: A descriptive and comparative study—I. Y. MAITEN
3. Graphic representation of a man by four-year-old children in nine prescribed drawing situations—P. F. CRIDLEY
4. Differences between two groups of adult criminals—R. S. TOLMAN
5. A comparative study by means of the Rorschach method of personality development in twenty pairs of identical twins—E. THOMP
6. Individual differences in the facial expressive behavior of preschool children: A study by the time-sampling method—C. SWAN

VOLUME 21—January-December, 1939

1. An experimental analysis of "level of aspiration"—R. GOULD
2. Some light on the problem of bilingualism as found from a study of the progress in mastery of English among preschool children of non-American ancestry in Hawaii—M. E. SMITH
3. Domination and social integration in the behavior of kindergarten children and teachers—H. H. ANDERSON
4. The capacity of the rhesus and cebus monkey and the gibbon to acquire differential response to complex visual stimuli—W. E. GALT
5. The social-sex development of children—E. H. CAMPBELL

VOLUME 22—January-December, 1940

1. Measuring human relations: An introduction to the study of the interaction of individuals—E. D. CHAPPELLE
2. Aggressive behavior in young children and children's attitudes toward aggression—M. D. FITE
3. Student attitudes toward religion—E. NELSON
4. The prediction of the outcome-on-furlough of dementia praecox patients—J. S. JACOB
5. Significant characteristics of preschool children as located in the Conrad inventory—K. H. READ
6. Learning by children at noon-meal in a nursery school: Ten "good" eaters and ten "poor" eaters—J. B. McCAY, E. B. WARING, AND P. J. KRUSE
7. Studies in the interpretation of play: I. Clinical observation of play disruption in young children—E. H. ERIKSON

VOLUME 23—January-June, 1941

1. An analysis of certain variables in a developmental study of language—F. M. YOUNG
2. Infant development under conditions of restricted practice and of minimum social stimulation—W. DENNIS
3. An analysis of the mental factors of various age groups from nine to sixty—B. BALINSKY
4. Factors influencing performance on group and individual tests of intelligence: I. Rate of work—M. W. BENNETT
5. Individual differences in apperceptive reaction: A study of the response of preschool children to pictures—E. W. AMEN

VOLUME 24—July-December, 1941

1. Twins T and C from infancy to adolescence: A biogenetic study of individual differences by the method of co-twin control—A. GESSELL AND H. THOMPSON
2. Finger nail-biting: Its incipency, incidence, and amelioration—A. L. BILLIG
3. An experimental study of the factors of maturation and practice in the behavioral development of the embryo of the frog, *Rana pipiens*—A. FROMME
4. The Fels child behavior scales—T. W. RICHARDS AND M. P. SIMONS
5. Measurement of the size of general English vocabulary through the elementary grades and high school—M. K. SMITH
6. Stereotypes in the field of musical eminence—P. R. FARNSWORTH

VOLUME 25—January-June, 1942

1. A study of factors determining family size in a selected professional group—J. C. FLANAGAN
2. A genetic study of geometrical-optical illusions—A. WALTERS
3. Interpretation of behavior-ratings in terms of favorable and unfavorable deviations: A study of scores from the Read-Conrad Behavior Inventory—K. H. READ AND H. S. CONRAD
4. Are there any innate behavior tendencies?—J. B. SCHOELLAND
5. An investigation of the intelligibility of the speech of the deaf—C. V. HADDINS AND F. C. NUMBERS



# Genetic Psychology Monographs (continued)

VOLUME 26—July-December, 1942

1. The critical frequency limen for visual flicker in children between the ages of 6 and 18—V. L. MILLER
- Some factors determining handedness in the white rat—K. L. WENTWORTH
2. Motivation and behavior—E. FRENKEL-BRUNSWIK

VOLUME 27—January-June, 1943

1. Comparison of children's personality traits, attitudes, and intelligence with parental occupation—N. R. MADDY
2. A comparative study of mental functioning patterns of problem and non-problem children seven, eight, and nine years of age—M. L. PIGNATELLI

VOLUME 28—July-December, 1943

1. Separation anxiety in young children: A study of hospital cases—H. EDLSTON
2. Correlates of vocational preferences—W. A. BRADLEY, JR.

VOLUME 29—January-June, 1944

1. Mental changes after bilateral prefrontal lobotomy—S. D. PORTEUS AND R. D. KEFNER
2. A twin-controlled experiment on the learning of auxiliary languages—B. PRICE, W. J. KOSTIR, AND W. M. TAYLOR

VOLUME 30—July-December, 1944

1. A method of administering and evaluating the thematic apperception test in group situations—R. M. CLARK
2. A study of anxiety reactions in young children by means of a projective technique—R. TEMPLE AND E. W. AMEN

VOLUME 31—January-June, 1945

1. The evolution of intelligent behavior in rhesus monkeys—B. WEINSTEIN
2. Perceptual behavior of brain-injured, mentally defective children: An experimental study by means of the Rorschach technique—H. WERNER

VOLUME 32—July-December, 1945

1. A clinical study of sentiments: I.—H. A. MURRAY AND C. D. MORGAN
2. A clinical study of sentiments: II.—H. A. MURRAY AND C. D. MORGAN

VOLUME 33—January-June, 1946

1. Interpretation of spontaneous drawings and paintings—T. S. WAERNER
2. Outstanding traits: In a selected college group, with some reference to career interests and war records—F. L. WELLS AND W. L. WOODS

VOLUME 34—July-December, 1946

1. The relation of emotional adjustment to intellectual function—J. L. DESPERT AND H. O. PIERCE
- The smiling response: A contribution to the ontogenesis of social relations—R. A. SPITZ
2. Finger-painting and personality diagnosis—P. J. NAPOLI

VOLUME 35—January-June, 1947

1. The thematic apperception technique in the study of culture-personality relations—W. E. HENRY
2. A continuation study of anxiety reactions in young children by means of a projective technique—M. DORKEY AND E. W. AMEN

A study of the vocational interest trends of secondary school and college women—A. M. CAWLEY

VOLUME 36—July-December, 1947

1. Maze test validation and psychosurgery—S. D. PORTEUS AND H. N. PETERS
2. The diagnostic implications of Rorschach's test in case studies of mental defectives—I. JOLLES

VOLUME 37—January-June, 1948

1. The radio day time serial: A symbolic analysis—W. L. WARNER AND W. E. HENRY
- The relation of personality characteristics and response to verbal approval in a learning task—G. L. GRACE
2. The mechanism of vision: XVIII. Effects of destroying the visual "associative areas" of the monkey—K. S. LASHLEY

A study of the relationship between handwriting and personality variables—P. CASTELNUOVA-TEDESCO

VOLUME 38—July-December, 1948

1. Modern language learning: The intensive course as sponsored by the United States Army, and implications for the undergraduate course of study—M. LIND
- Conflict: A study of some interactions between appetite and aversion in the white rat—M. A. TOLCOTT
2. Schizophrenia and the MAPS test: A study of certain formal psycho-social aspects of fantasy production in schizophrenia as revealed by performance on the Make a Picture Story (MAPS) Test—E. S. SHNEIDMAN

A study of the transmission of authority patterns in the family—H. L. INCERSOLL

VOLUME 39—January-June, 1949

1. A study of the psychoanalytic theory of psychosexual development—G. S. BLUM
- The assessment of parental attitudes in relation to child adjustment—E. J. SHOBEN, JR.
2. Qualitative differences in the vocabulary responses of normals and abnormal—H. FELTEL

The relative effectiveness of motion and still pictures as stimuli for eliciting fantasy stories about adolescent-parent relationships—P. E. EISNER

The organization of hereditary maze-brightness and maze-dullness—L. V. SEARLE

VOLUME 40—July-December, 1949

1. An experimental study of what young school children expect from their teachers—B. BIBER AND C. LEWIS
- A study of the relative effects of age and of test difficulty upon factor patterns—H. A. CURTIS
2. A projective experiment using incomplete stories with multiple choice endings—J. K. SEATON

Effects of sex role and social status on the early adolescent personality—E. MILNER

Social perceptions and attitudes of children—M. RADKE, H. TRAGER, AND H. DAVIS

VOLUME 41—January-June, 1950

1. Some psychological and educational aspects of pediatric practice: A study of well-baby clinics—L. H. BLUM
- One-trial learning in the domestic rat—B. B. HUDSON
2. An introduction to the principles of scientific psychoanalysis—A. ELLIS

Awareness of racial differences by preschool children in Hawaii—D. V. SPRINGER

Age trends in children's evaluation of teacher-approved and teacher-disapproved behavior—S. L. WITRYOL

The relationship between level of vocational aspiration and certain personal data: A study of some traits and influences bearing on the prestige level of vocational choice—J. STUBBINS

VOLUME 42—July-December, 1950

1. Personality patterns of suicidal mental hospital patients—N. L. FARDROW
- Sex-role identification in young children in two diverse social groups—M. RABBAN

A study of the influence of the social field on individual behavior: As revealed in the expression of hostility and warmth by neurotics and paranoid schizophrenics in discussion group situations—D. SHAPIRO

An experimental study of avoidance—R. F. HEFFERLINE

VOLUME 43—January-June, 1951

1. A study of copying ability in children—E. A. TOWNSEND
- Prestige motivation of gifted children—D. P. ACUSUBEL
2. A psychological study of physical scientists—A. ROE

VOLUME 44—July-December, 1951

1. The organization of hostility controls in various personality structures—S. FISHER AND E. HINDS
- Children and radio: A study of listeners and non-listeners to various types of radio programs in terms of selected ability, attitude, and behavior measures—E. A. RICCIUTI

Quantitative expression in young children—W. E. MARTIN

The use of magnetic devices in the collection and analysis of the preverbal utterances of an infant—A. W. LYNIP

VOLUME 45—January-June, 1952

1. Japanese-American personality and acculturation—W. CAUBILL
2. A statistical study of the Freudian theory of levels of psychosexual development—C. A. BARNES

Personality characteristics of selected disability groups—D. N. WIENER



# Genetic Psychology Monographs (continued)

VOLUME 46—July-December, 1952

1. The relationship of social status, intelligence, and sex of ten- and eleven-year-old children to an awareness of poverty—F. J. ESTYAN  
An empirical study of the castration and Oedipus complexes—S. M. FRIEDMAN
2. The relationship between projective test scoring categories and activity preferences—M. M. SCHWARTZ  
A comparison of formal and content factors in the diagnostic testing of schizophrenia—M. SHERMAN

VOLUME 47—January-June, 1953

1. Ability and accomplishment of persons earlier judged mentally deficient—D. C. CHARLES  
Variations in the consistency of the behavioral meaning of personality test scores—M. KORNREICH
2. Some child-rearing antecedents of aggression and dependency in young children—R. R. SEARS, *et al.*  
Symptom correlates for descriptive diagnosis—J. R. WITTENBORN, *et al.*

VOLUME 48—July-December, 1953

1. Age and mental abilities: A longitudinal study—W. A. OWENS, JR.  
The development of a personality questionnaire for drinkers—P. J. HAMPTON  
Personality and physical disease: A test of the Dunbar hypothesis applied to diabetes mellitus and rheumatic fever—D. H. CHOWELL
2. Socio-economic contrasts in children's peer culture prestige values—B. POPE  
A critical review of the stability of social acceptability scores obtained with the partial-rank-order and the paired-comparison scales—S. A. WITRYLO AND G. G. THOMPSON  
A study of the effects of color on Rorschach responses—G. G. BRODY

VOLUME 49—January-June, 1954

1. Factors underlying major reading disabilities at the college level—J. A. HOLMES  
Parent behavior toward first and second children—J. K. LASKO
2. Social-status and intelligence: An experimental study of certain cultural determinants of measured intelligence—E. A. Haggard  
Certain determinants and correlates of authoritarianism—S. SIEGEL  
Personalities in faces: I. An experiment in social perceiving—P. F. SECORD, W. F. DUKES, AND W. BEVAN

VOLUME 50—July-December, 1954

1. A study of the relationship between play patterns and anxiety in young children—E. W. AMEN AND N. RENISON  
Operational exploration of the conceptual self system and of the interaction between frames of reference—M. EDELSON AND A. E. JONES  
Problem solving: A statistical description of some relationships between organismic factors and selected response measures—N. A. FATTU, E. KAPOS, AND E. V. MECH
2. The relation of cortical potentials to perceptual functions—C. CHAYATTE  
The import for clinical psychology of the use of tests derived from theories about infantile sexuality and adult character—D. W. MILES  
Measuring personality in developmental terms: The Personal Preference Scale—M. H. KROUT AND J. K. TABIN

VOLUME 51—January-June, 1955

1. Some relations between techniques of feeding and training during infancy and certain behavior in childhood—A. BENNETTELM  
The expression of personality in drawings and paintings—L. H. STEWART
2. Negative stereotypes concerning Americans among American-born children receiving various types of minority-group education—J. A. FISHMAN  
The Lincoln-Oseretsky Motor Development Scale—W. SLOAN

VOLUME 52—July-December, 1955

1. Some personality correlates of sex, sibling position, and sex of sibling among five- and six-year-old children—H. L. KOCH  
A quantitative Rorschach assessment of maladjustment and rigidity in acculturating Japanese Americans—G. DeVos  
Measurement of authoritarianism and its relation to teachers' classroom behavior—H. M. MCGEE
2. The formal aspects of schizophrenic verbal communication—B. MIMM  
A study in an aspect of concept formation, with subnormal, average, and superior adolescents—H. N. HOFFMAN  
Traumatic avoidance learning: Acquisition and extinction in dogs deprived of normal peripheral autonomic function—L. C. WYNNE AND R. L. SOLOMON

VOLUME 53—January-June, 1956

1. As the psychiatric aide sees his work and problems—F. L. WELLS, M. GREENBLATT, AND R. W. HYDE  
An investigation of avoidance, anxiety, and escape behavior in human subjects as measured by action potentials in muscle—J. D. BROTHERS  
Spread of effect: A critical review—M. H. MARX
2. Stress, fantasy, and schizophrenia: A study of the adaptive processes—O. J. B. KERNER  
The attitude structure of the individual: A Q-study of the educational attitudes of professors and laymen—F. N. KERLINGER

VOLUME 54—July-December, 1956

1. A study of personality differences between middle and lower class adolescents: The Szondi Test in culture-personality research—L. RAINWATER  
The assessment of parental identification—S. W. GRAY AND R. KLAUS
2. The influence of social context on impulse and control tendencies in preadolescents—G. H. ZUK  
Tender-mindedness versus tough-mindedness in psychology: A reexamination—H. WINTROP  
A method for the comparison of groups: A study in thematic apperception—L. C. SCHAW AND W. E. HENRY

VOLUME 55—January-June, 1957

1. Academic performance and personality adjustments of highly intelligent college students—B. M. HORRALL  
The use of the Vineland Social Maturity Scale in the planning of an educational program for non-institutionalized low-grade mentally deficient children—M. G. GOTTSCHEN
2. The structure and origin of the anal character—H. BELOFF  
Free expression of adolescents' interests—M. AMATORA  
The rôle of mass media and the effect of aggressive film content upon children's aggressive responses and identification choices—R. S. ALBERT  
Interest in persons as an aspect of sex difference in the early years—E. W. GOODENOUGH

VOLUME 56—July-December, 1957

1. Some stable response determinants of perception, thinking, and learning: A study based on the analysis of a single test—M. L. SIMMEL AND S. COUNTS  
Potentials of age: An exploratory field study—S. L. PRESSEY
2. Tactile communication—L. K. FRANK  
The use of a filmed puppet show as a group projective technique for children—M. R. HAWORTH  
The social competence of middle-aged people—R. J. HAVIGHURST

VOLUME 57—January-June, 1958

1. Psychological and cultural problems in mental subnormality: A review of research—S. B. SARASON AND T. GLADWIN  
Developmental aspects of discrimination in relation to adjustment—P. LONDON  
Muscular tension as a response to psychological stress in rheumatoid arthritis and peptic ulcer—J. A. SOUTHWORTH  
Frustration: Theory and experiment—R. LAWSON AND M. H. MARX

VOLUME 58—July-December, 1958

1. Emotional aspects of political behavior: The woman voter—E. M. BENNETT AND H. M. GOODWIN  
The accuracy of self estimate: A measure of self-concept reality—R. M. BRANDY  
Personality factors in social mobility: A study of occupationally mobile businessmen—J. C. ABREGLEN
2. Interaction between authoritarian and nonauthoritarian principals and teachers—P. LAMBERT  
Some factors related to the choice-status of ninety eighth-grade children in a school society—D. ELKINS  
Some relationships between fantasy and self-report behavior—R. C. CALOGERAS



# Genetic Psychology Monographs (continued)

VOLUME 59--January-June, 1959

1. Freshman Rorschach and college performance—G. R. SCHMEIDLER, M. J. NELSON, AND M. BRISTOL  
Perceptual aspects of attributed movement—R. W. COAN  
Men and women: Personality patterns and contrasts—E. M. BENNETT AND L. R. COHEN
2. Personality factors in mothers of cerebral palsied children—G. BOLES  
Separation and integration: A social psychological analysis of editorial content in New York newspapers of three American minority groups—J. A. FISHMAN AND C. S. FISHMAN  
Self, role, and satisfaction—A. L. BORNH

VOLUME 60--July-December, 1959

1. The constancy of personality ratings over two decades—R. D. TUDDENHAM  
Mother-son identification and vocational interest—L. H. STEWART  
A study of mother-child relationships in the emotional disorders of children—M. J. ROSENTHAL, M. FINKELSTEIN, E. NI, AND R. E. ROBERTSON  
On the trail of the wolf-children—W. F. OGURN AND N. K. BOSE
2. Measuring the mental health of normal adults—R. F. PECK  
Parental identification in young children—W. EMMERICH  
The problems of Jewish religious youth—B. M. LEVINSON

VOLUME 61--January-June, 1960

1. Identifying socially maladjusted school children—A. A. ADAMS  
Social interaction in an interracial nursery school—H. W. STEVENSON AND N. G. STEVENSON, *et al.*  
Interests of pre-adolescent boys and girls—M. AMATURA  
Educational and sex differences in the organization of abilities in technical and academic students in Columbia, South America—J. F. FIDELLA
2. Some themes in the personalities of German men—L. RAINWATER  
Right-left gradients in body image, body reactivity, and perception—S. FISHER  
Longitudinal survey of child Rorschach responses: Younger subjects two to 10 years—L. B. AMES  
Personality patterns related to occupational roles—M. SIEGELMAN AND R. F. PECK

VOLUME 62--July-December, 1960

1. The reading process: A descriptive, interdisciplinary approach—M. P. LANGMAN  
Reactions to stress as a function of the level of intelligence—S. M. H. ZAIDI  
Rorschach diagnosis by a systematic combining of content, thought process, and determinant scales—P. A. BOWER, R. TESTIN, AND A. ROBERTS  
Longitudinal survey of child Rorschach responses: Older subjects aged 10 to 16 years—J. B. AMES
2. Twenty years of shock therapy in America, 1937-1956: An annotated bibliography—N. H. FRONKO, R. NITTERLY, AND K. BERG  
The quantitative analysis of parent behavior toward psychotic children and their siblings—L. M. DONNELLY

VOLUME 63--January-June, 1961

1. Make a sentence test: An approach to objective scoring of sentence completions—E. F. BORLATT  
Some aspects of the relationship between perception and motility in children—E. P. ROTHMAN  
Some psychophysical aspects of mental illness in children—P. H. BARKOWITZ  
Conceptual thinking in schizophrenic children—G. FRIEDMAN
2. Some maternal influences on children's personality and character—J. C. FINNEY  
A study of concept formation in normals, mental defectives, and brain-damaged adults—R. W. ZASLOW  
An investigation of trait relationships among six-year-old children—S. A. AVAKIAN

VOLUME 64--July-December, 1961

1. Coping patterns of preschool children in response to intelligence test demands—A. MORRIS  
Studies of individual and paired interactional problem-solving behavior of rats: II. Solitary and social controls—A. BARON AND E. A. LITTMAN
2. Studies of role performance—J. H. MANN  
A psychological assessment of professional actors and related professions—R. TART  
A cross-cultural study of menstrual taboos—W. N. STEPHENS  
The biological roots of creativity—H. GUTMAN

VOLUME 65--January-June, 1962

1. A Rorschach study of the development of personality structure in white and Negro children in a Southeastern community—A. C. PRICE  
Measuring personality patterns of women—J. LUEVINGER  
Home interests in early adolescence—M. AMATURA  
Social and emotional adjustment during adolescence as related to the development of psychosomatic illness in adulthood—L. H. STEWART
2. A systematic study of interaction process scores, peer and self-assessments, personality and other variables—E. F. BORLATT  
Children's conceptions of the content and functions of the human body—E. GELBERT

VOLUME 66--July-December, 1962

1. Ocular-manual laterality and perceptual rotation of literal symbols—J. FLESCHER  
Parental child-rearing attitudes and their correlates in adolescent hostility—S. B. CHODURA  
Some explorations into the nature of anxieties relating to illness—P. R. KOCHINS  
Developmental-interactive aspects of hypnosis. Some illustrative cases—J. R. HILGARD AND E. R. HILGARD
2. Teaching a two-year-old to read: An experiment in early childhood learning—W. FOWLER  
Achievement gains or losses during the academic year and over the summer vacation period. A study of trends in achievement by sex and grade level among students of average intelligence—K. M. PANSLEY, JR., AND M. POWELL

VOLUME 67--January-June, 1963

1. Verbatim dialogue of a mother and child in therapy—C. E. MOUSTAKAS  
A study of self-concept: Delinquent boys' accuracy in selected self-evaluations—W. E. AMOS  
An integrating study of the factor structure of adult attitude interests—R. B. CATTELL AND J. HORN  
Social factors in mental disorders: A contribution to the etiology of mental illness—W. G. COLEMAN
2. Conceptual preferences associated with choice of quantitative and nonquantitative college majors—J. GILBERT  
A study of the concept of the feminine role of 51 middle-class American families—A. STEINMANN

VOLUME 68--July-December, 1963

1. Brain dynamism as reflected in illusions—G. K. YACOWYNSKI  
The nature and measurement of components of motivation—R. B. CATTELL, J. A. RADCLIFFE, AND A. B. SWENEY
2. Sharing in preschool children as a function of amount and type of reinforcement—W. F. FISCHER  
The Gesell Incomplete Man Test as a measure of developmental status—L. B. AMES AND F. L. ILG  
Psychological and sociological factors in prediction of career patterns of women—M. C. MULLVY  
Functionalism and its historical significance—R. HARRISON

## PUBLICATION IN THE JOURNAL OF PSYCHOLOGY

*The Journal of Psychology* provides a unique service to authors. Articles accepted for publication in this periodical appear in print with the least possible amount of delay. In line with this policy, all steps of the acceptance procedure are maximally streamlined, and the author's work is presented with a bare minimum of editing or restriction of any kind. *The Journal of Psychology* is the ideal medium for presentation of timely professional material.

Authors are requested to exercise unusual care in the preparation of papers for this journal. Accuracy, grammar, and style should be checked and rechecked before a clean version of the manuscript is submitted. Careful attention should be paid to the General Instructions on the back cover of each issue and to the announcement on the inside front cover.

Authors are particularly requested to note the fact that all parts of the manuscript—including references, footnotes, and tables—must be double-spaced.

## ERRATA

PHELAN, JOSEPH G., & GUSTAFSON, CHARLES W. Reversal and nonreversal shifts in acute brain-injured with injury diffusely localized. *J. of Psychol.*, 1968, 70, 249-259. The word "cerebrovascular" should be substituted throughout the paper for the word "cardiovascular."

\$30.00 per annum  
\$15.00 per volume  
\$7.50 per half volume

**BIMONTHLY**  
Three volumes per year  
Immediate publication

January, 1969  
Volume 71, First Half

(Founded in 1935 by Carl Murchison)

# The Journal of PSYCHOLOGY

The General Field of Psychology

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold

under the restrictions of a bilateral contract, and may not be resold for a period of five years from the date of publication.

JANUARY, 1969

(Manuscripts are printed in the order of final acceptance)

Notes on photographic memory . . . . .	3
BY NORMA E. CUTTS AND NICHOLAS MOSELEY	
Abasement scores and adjustment of neuropsychiatric patients . . . . .	17
BY ALLEN GOSS	
Psychodynamic factors in the experimental investigation of hypnotically induced emotions with particular reference to blood glucose measurements . . . . .	21
BY MILTON V. KLINE AND MORTON LINDER	
The use of figure drawings to assess religious values . . . . .	27
BY CLIFFORD ATKISSON, LEONARD HANDLER, AND RAYMOND R. SHRADER	
Semantic space is (approximately) bipolar . . . . .	33
BY P. M. BENTLER	
Historical and contemporary attitudes toward numbers . . . . .	41
BY M. S. LINDAUER	

(OVER)

Copyright © 1969 by The Journal Press  
Provincetown, Massachusetts, U. S. A.

All rights reserved

Published at 35 New St., Worcester, Mass.

Second-class postage paid at Worcester, Mass.

Send all mail to Managing Editor, 2 Commercial Street, Provincetown, Massachusetts 02657



A comparison of personality factor structure in British and American university students . . . . .	45
BY KAY JAMISON AND ANDREW L. COMREY	
Relation between internal-external control and aggression . . . . .	59
BY CHRISTENE BLANTON WILLIAMS AND FLORENCE E. VANTRESS	
Experimenter expectancy effects in three dissimilar tasks . . . . .	63
BY RICHARD L. WESSLER	
Retroactive inhibition of connected discourse as a function of meaningfulness and of degree of original learning . . . . .	69
BY JOHN A. MILLS AND GORDON WINOCUR	
Performance characteristics of field dependent and independent individuals on an auditory signal detection task . . . . .	77
BY VICTOR J. DEFazio AND WILLIAM F. MORONEY	
The effect of leading contour on the relative lengths of moving light and dark arcs . . . . .	83
BY GORDON STANLEY AND ROBERT JACKSON	
An investigation of some intrinsic variables affecting test responses . . . . .	89
BY JOHN BOTTRILL	
S-R and R-S paired-associate learning as a function of concreteness, imagery, specificity, and association value . . . . .	95
BY A. DANIEL YARMEY AND BRIAN J. O'NEILL	
Visual confusion matrices: Fact or artifact? . . . . .	111
BY DENNIS F. FISHER, RICHARD A. MONTY, AND SAM GLUCKSBERG	
Further study of performance errors on Ravens Progressive Matrices (1938) . . . . .	127
BY CHARLES SHEPPARD, DIANE FIORENTINO, LOIS COLLINS, AND SIDNEY MERLIS	
Comparison between American permanent residents of Israel: Part I, American background . . . . .	133
BY GERALD ENGEL	
Problems in the cross-cultural study of schizophrenia . . . . .	143
BY IHSAN AL-ISSA	

## NOTES ON PHOTOGRAPHIC MEMORY\*

*Branford, Connecticut; and Sandy Hook, Connecticut*

NORMA E. CUTTS AND NICHOLAS MOSELEY

### A. INTRODUCTION

It has been said that there are people who can reel off any number of verses which they have heard only once, but I never happened to be present when this was done. However, one must trust it can be done, because only he who believes may hope.

Quintilian (*cf.*, 40)

"I wish I had a better memory" is an almost universal aspiration. Certainly a good memory is a valuable asset. How valuable it may be is indicated by an analysis, hitherto unreported, which the authors have made of the abilities of 150 eminent men and women as these are portrayed in their biographies. The authors find 58 whose memory is described in some detail. Of these, 24 are described by their biographers as having what amounts to "photographic memory," 14 extraordinary, 13 excellent, and 7 average or poor. Actually, of the last group, three who rated themselves as having poor memories—Gable (9), Maugham (37), Wells (50)—<sup>1</sup> had, according to their own accounts of their work, superior memory. One—Darwin (19)—compensated for an average memory by keeping voluminous notes arranged systematically. Two were athletes who drilled themselves indefatigably in the skills of their sports—Dempsey (15), Ruth (14).

The large proportion of eminent people who have superior memories accords with the findings of Miles's study of traits of genius, where "strength of memory" is the trait which correlates most closely with estimated intelligence (38). It also supports Wechsler's conclusion that "undifferentiated memory" is a separate factor in intelligence (48).

The authors were particularly struck by the number of their subjects with "photographic memory."<sup>2</sup> This is discussed in psychology texts under "eidetic

\* Received in the Editorial Office, Provincetown, Massachusetts, on September 4, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> All persons named between dashes are the subjects of biographies cited in the References.

<sup>2</sup> Balzac (38), Barbour (49), Buchan (3), Curie (17), Davis (5), de Gaulle (26), Edison (32), Freud (31), Frost (16), Fuller (22), Hadley (25), Hoover (35), Kennedy (7), Lewis (44), Liebig (47), Lovett (interview), Maazel (39), Macaulay (46), Mayo, Charles (13), Michelson (30), Roosevelt (8, 41), Toscanini (10), Waksman (47), Wright, Wilbur (36), Yardley (52).

imagery" and is described as the ability to look briefly at materials, including pictures and printed matter, and to recall them as visual images. The image is said to be very exact in detail, if not always eidetic: i.e., a complete reproduction like a photographic print [*cf.*, Ruch (42), Allport (1), Feldman (20)]. The term is popularly used to describe any extraordinary memory which enables a person to apprehend materials of any kind with great speed and to reproduce them with unusual accuracy after long periods of time. This popular usage is supported by the biographies.<sup>3</sup> These portray individuals whose extraordinary memories differ according to the principal means of perception (visual, auditory, olfactory, tactile) and as to types of materials (verbal, mathematical, spatial).

Even granting the broad definition, the present authors had never in all their combined experience of more than 100 years as teachers been aware of having a student with the extraordinary speed of learning and accuracy of recall which some of the biographies describe. The present authors had known personally only two persons with these powers, the late Arthur T. Hadley, former President of Yale, and Robert A. Lovett, former Secretary of Defense. In order to discover whether the phenomenon was as rare as the experience of the authors indicated, they tested more than 1000 elementary- and secondary-school students and more than 100 adults. One test consisted of exposing a poster-size picture for 30 seconds and asking the subjects to answer questions on details: e.g., how many windows in the picture of a house. Another required the repetition of a 100-word prose passage read only once. Another required the repetition of a similar passage read aloud to the subjects, and another the repetition of a similar passage which they copied out for themselves reading aloud as they did so.

The students who did best on the tests and all of the adults were interviewed and questioned about their visual and auditory mental processes as they tried to learn the materials and while they tried to recall them.

The results of these tests and interviews are reported here for the first time.

The authors' screening and the interviews do not disclose any individual with a memory as great as Hadley's or Lovett's or those described in some of the biographies, notably those of Balzac (38), Macaulay (46), and Toscanini (10). However, they do confirm a close relation of excellent memory to high achievement in school, college, and careers. In fact, the students who ranked in the upper quartile on the memory tests invariably ranked in the upper quartile in scholastic standing. (The reverse, of course, does not hold. Effort and persistence may compensate for relatively poor powers of memory.)

<sup>3</sup> Throughout this article, "the biographies" refers to those cited in the References.



The tests and interviews also show that practically everyone uses some visual imagery, but that there are great individual differences. Some people, for example, report that to remember visual details they must describe them to themselves verbally and that they recall the words when trying to image even a picture or a spatial figure. When visualizing a picture or a page, most tend to place it, to "project" it, in the position where it was when they were learning, but a few place it a few inches above and to the right or left of their foreheads. Some just say, "It is there" [*cf.*, Allport (1), Feldman (20)].

The people with superior memories whom the authors tested are generally, but not always, fast readers. The fastest readers incline to skim, picking out words keyed to their special interests. When recalling, they can give only the gist of a paragraph. When asked to learn a passage in order to repeat it verbatim, they tend to go over it again and again and may actually move their lips as they do so. In contrast, some of the present authors' cases from the biographies not only read very rapidly, but also recall much of what they have read verbatim—Balzac (38), Curie (17), Hadley (25), Macaulay (46).

The most complete and accurate immediate recall of printed materials by the adults and students interviewed was for passages which they copied out, reading them aloud as they did so: i.e., when they perceived the words visually, orally, and kinaesthetically.

Many people, including some who have only average memory, have vivid visual recall of long-past situations that they found emotionally moving, including situations which caused them sorrow, fear, or embarrassment [*cf.*, Darwin (19)].

The general conclusion of the present writers is that there is no such thing as photographic memory in the literal sense of taking a snapshot of a page and filing it in the mind like a photographic print which can be examined at will [*cf.*, Hunter (29), Kimble (34)]. In the assessment of memory the authors think the best procedure is to describe it in general terms with appropriate qualifications: e.g., an excellent visual memory, especially for figures. Since there is so much interest in extraordinary memory and so much misunderstanding even among psychologists, it seems to the authors worthwhile to report the two cases of which they have personal knowledge and, using these and the materials which they have collected from biographies, interviews, and tests, to discuss some of the questions most often raised.

#### B. ARTHUR T. HADLEY

[This account draws on personal associations extending over 15 years, on discussions with President Hadley's children, and on the biography written by his son (25)].

Arthur Twining Hadley, 1856-1930, was the son of James Hadley, 1821-1872, Professor of Greek in Yale College.

James Hadley had a powerful memory, as evidenced by his knowledge of Latin, Greek, Hebrew, German, Spanish, Italian, Sanskrit, Gothic, Welsh, and Gaelic among other languages and by his habit of reporting in his diary summaries of sermons that he had heard (25).

His son, Arthur, graduated from Yale at 20 as valedictorian of his class and took his Ph.D. at Berlin. His special fields were labor relations, statistics, and economics with particular reference to railroad transportation. His biographer (25, p. 56) says that Arthur Hadley felt "he had been fortunate in selecting a subject which was of real importance but where the literature, although growing fast, did not yet exceed in volume what one man could read. He believed that he had studied everything written on the subject, not only in book form, which was an easy task, but in periodicals and technical reports and statistical sources, in foreign languages as well as in English. In this he was aided by the fact that, partly by nature and partly by training, he was a phenomenally rapid and retentive reader. He could go through a book, turning the leaves so rapidly that he would appear to be merely skimming over them, and at the end not only retain the gist of the book but be able to quote, or at least paraphrase, long passages from it" [*cf.*, Balzac (38), Curie (17), Macaulay (46)].

At the celebration of Yale's bicentennial, the delegate from the University of Upsala spoke in Latin. The alumni were delighted when Hadley replied in Latin. One of the present writers asked him about this years later. He said that, as soon as the man from Upsala started to speak, he realized he must reply in Latin. "Fortunately," Hadley said, "the vocabulary of compliment in any language is small, so as he made his remarks I coiled them down, and when the time came rose and paid them out in reverse order."

Hadley was famous for his memory of vintages. At a dinner party when he was Visiting Professor in Berlin he was challenged to name a wine. "He recognized the general district and the year without difficulty, but after several sips confessed that he could not name the vineyard . . . [though he found the wine] markedly reminiscent of three well-known vineyards, which he named. The host and guests listened with mounting excitement. . . . Said the host . . . 'You do not realize what you have done. You could never have tasted this wine. It is from a small private vineyard that has been in my family for some centuries and the wine never comes on the market. The three vineyards which you named, however, are the three by which my family's vineyard is bounded' " (25, p. 186).

When he was 70, one of the present writers spoke to him in praise of a pas-

sage of Lucian. He looked at the ceiling and repeated the page word for word in Greek. Asked when he had last read it, he replied, "When I was a sophomore in college"—i.e., 52 years before. He always lectured and made speeches without notes, although he kept three phrases of the Lord's Prayer on the chapel lectern because he feared he might get the order wrong. When he knew he was to make a speech he composed it mentally, unconsciously making the appropriate gestures as he did so. (Once in Yale Chapel, sitting on the platform behind a rather dull visiting preacher, he planned a speech to be made at a fraternity house that evening. His gestures, his facial expression, his bows to imaginary applause delighted the students, horrified his family, and caused the unseeing preacher to remark afterwards that the students seemed unusually attentive.)

Hadley's ability to quote accurately long poems or prose passages or elaborate statistics remained throughout his life. Though none of his children, grandchildren, or great-grandchildren has a memory equal to his, most of the children and grandchildren have excellent memories and have had distinguished records in college and in their professions.

#### C. ROBERT A. LOVETT

Robert A. Lovett was born in 1895. He prepared for Yale at the Hill School, graduated from Yale in 1918, and attended the Harvard Law School in 1919-1920 and the Harvard School of Business Administration in 1920-1921. He was a naval pilot in 1917-1918, advancing to lieutenant commander and winning the Naval Cross. He has been a partner in Brown Brothers Harriman & Co., private bankers, since 1926, except when in government service. In World War II he was Assistant Secretary of War for Air, and after the war Undersecretary of State, 1947-1949, Deputy Secretary of Defense, 1950-1951, and Secretary of Defense, 1951-1953 (51). The following account is based on an interview he gave the present writers in August 1966.

Lovett's father educated himself as a lawyer, became counsel for E. H. Harriman of the Union Pacific Railroad, and later president of the railroad. Lovett's mother was a graduate of Wellesley and taught at Texas Peabody. When Lovett was going to school in New York, his father would question him every evening on what he had seen en route and insist on an accurate report of details. When his father judged Lovett's observation accurate and complete in detail, he rewarded him with 25 cents; when inaccurate or incomplete he docked him 25 cents. It was not enough to say he had seen a dray loaded with steel; he had to describe the horses, the harness, the number of beams, and the way they were fastened by steel chains.

Lovett was first aware that he had an unusual memory when a Greek mas-



ter at the Hill School seemed amused by his ability to master paradigms of verbs. He excelled in French, Greek, and Latin in school, and in mathematics and languages in college. He found school and college courses "a snap." He was a member of Phi Beta Kappa at Yale.

Some anecdotes told by Lovett illustrate his memory. In flying submarine patrol he visualized the details of the chart, last reported sightings of submarines, and positions of surface craft. Life depended on alert observation and accurate imagery.

At Harvard Law School he was taking a course in torts with Dean Roscoe Pound. He was also engaged in a moot court and had worked hard studying pertinent cases. On his final examination there was a question about one of these cases, which he answered by quoting it verbatim, complete with case numbers of the references. His paper was returned with "Please see me. R.P." scrawled on it. When he arrived at the Dean's office, Pound, wearing a green eyeshade, was sitting behind his desk, obviously embarrassed. Finally he said, "Mr. Lovett, you realize I must give you either an A or a zero." "Would it help, Dean Pound, if I quoted the case to you now?" Lovett proceeded to do so. Pound understood because he himself had a "photographic" memory. Lovett can still quote the case and the references. [A similar anecdote is told, without names, by Floyd L. Ruch (42).]

When Lovett was starting in banking as a clerk in the National Bank of Commerce, one of his duties was to verify the checks and statements prepared each evening for the New York Clearing House. When there was a discrepancy in the totals, apparently due to a transposition, he could visualize the individual checks cashed by each teller and select the check which had been wrongly entered.

When he was Secretary of Defense, he regularly appeared before a Congressional committee to explain and justify the Army's budget. Other Cabinet members would appear backed by crowds of clerks carrying books and ledgers marked with slips of paper. Lovett appeared alone and was able to answer all questions by quoting the pertinent figures without using notes. Soon the Congressmen became so interested in the performance that they paid more attention to it than to the Budget! [*Cf.*, Macaulay's similar account of his experience as Secretary of War supporting the Army estimates before Parliament (46).]

Six weeks before the interview, Lovett said, he had had occasion to telephone Chicago. He had not thought of it since and had no special reason to remember the telephone number, but he quoted it to the present authors.

Under catechizing from the interviewers, Lovett stated that his memory

frequently seemed to operate involuntarily. After a drive on the crowded Long Island Expressway his children questioned him about the license numbers of cars that had passed them, and he recalled them without error. (The children had made notes.) After a dinner party he can describe to his wife the details of color and cut of the dresses worn by the other ladies present. When he has examined a painting with close attention, he recalls the colors perfectly—"like absolute pitch"—but is not as good at this as his wife.

Sometimes he wonders if his mind is not cluttered with useless information.

His memory is largely visual. He sees print (or a color or a picture) in "the form in which it occurred." He remembers what he has seen better than what he has heard. In the case of a telephone number or date he remembers it only if he has seen it in a book or written it down. He does not know German, but could remember German words if they were printed in roman type, though not if they were printed in German type.

His memory for numbers is better than for the printed word or for colors.

He has a very definite recall of odors.

He seems to remember some things permanently and to forget some. He remembers dates for months and then they may "taper off." Use, practice, and a sense of urgency increase the likelihood of permanence.

He is uncertain about the effect of alcohol. Ulcers kept him from drinking during World War II. Now, he says, alcohol certainly does not sharpen his memory, and perhaps prevents his caring.

Fatigue interferes with his memory.

He says that his memory has definitely deteriorated since he was 69. (We saw no evidence of deterioration.) He attributes this to declining eyesight. Until recently he has never had to use glasses. He now uses weak bifocals.

When visualizing a printed page he sees the beginning and end of a paragraph and "most in between."

He reads very rapidly.

No one else in his family has ever had a photographic memory so far as Lovett knows. His son and daughter had excellent scholastic records. His son ranked in the top three in his Yale class despite very poor vision; he is a very hard worker.

Lovett attributes his superior memory to the training in observation and recall which his father gave him as a boy. He thinks that, as in the case of patrolling pilots and of bank clerks, memory is improved by "use, practice, and a sense of urgency." He regrets that his governmental duties kept him from trying to train his own children to remember.

## D. DIFFERENCES IN SENSUAL PERCEPTION

The type of memory most frequently described by the writers' subjects is that for materials perceived visually [*cf.*, Gregory (24)]. Many of them learn whole paragraphs after one or two or three readings—Buchan (4), Curie (17), Freud (31), Macaulay (46), Yardley (52). Some, like Lovett (interview), must see materials which they wish to recall—Waksman (47), von Liebig (47). Some apparently have equal visual and auditory powers—Frost (16), Lewis (44), Maazel (39). If vision is defective, tactile and olfactory sensations become the basis of memory—Keller (33). A few individuals seem to have almost universal powers—visual, auditory, olfactory, tactile—Balzac (38), Hadley (25 and interviews), Toscanini (10). There is little evidence as to how much these differences are due to innate capacities; but there is evidence that clearly supports the need for close attention and intent to remember—Churchill (11), Curie (17). "Apparently inadvertent" learning—Lovett (interview), Macaulay (46)—seems to be the result of training that establishes a habit of close observation with the expectation of recall.

## E. DIFFERENCES RELATED TO TYPES OF MATERIALS

The old observation that some students are better at reading than at arithmetic, and *vice versa*, is borne out by the subjects of this study—Churchill (12), Roosevelt (8), Hilton (28), Hoover (35). There are cases where the ability to recall spatial materials surpasses the ability to recall other types of materials—Edison (32), Ford (6), Orville Wright (36). One individual with superior memory whom the authors tested by matched pairs of designs learned six pairs in 30 seconds and reproduced them as fast as she could draw. When questioned, she said she had given each design of each pair a name—e.g., a pentagon and a railroad track—recalled the names, and drew her representation of the name.

Some relative deficiencies may be explained by lack of training—Edison (32)—but some persist despite high general intelligence and prolonged efforts at improvement—Churchill (12), Hoover (35). As with differences related to the senses, there are some individuals who have extraordinary recall for all types of materials—Davis (5), Hadley (25 and interviews), Macaulay (46), Toscanini (10).

The possible differentiation of memory for materials associated with different factors in intelligence has obvious implications for theories of intelligence and for intelligence- and aptitude-testing.



## F. EXTENT AND DURATION OF RECALL

Several of the subjects of the biographies say that they could reproduce whole volumes of materials—e.g., the whole of *Pilgrim's Progress*—Buchan (3), Macaulay (46)—but none seems to have been put to the test. There are, however, instances of the more or less permanent retention of rather extensive materials of many kinds—Barbour (49), Churchill (11), Davis (5), Edison (32), Hadley (interviews), Lovett (interview), Macaulay (46), Toscanini (10). Purposeful learning of extensive materials involves both effort and time—Lincoln (43), Macaulay (46). Macaulay reports that it took him two hours to learn Act IV of *The Merchant of Venice* although he already knew 150 of the 400 lines (46). Several speak of "learning by heart," implying both intent and effort—Buchan (4), Churchill (11). Some individuals say that they retain so many facts so long that their minds are cluttered up. Macaulay had to avoid a tendency to quote interminably (46). Lovett speaks of "clutter" (interview). Balzac (38) had difficulty in selecting details pertinent to a story. However, their achievements belittle the difficulties.

## G. AGE

Allport (1) concluded from his studies of eidetic imagery in children that it began to decline in the early teens. This conclusion is often quoted in elementary texts of psychology and applied to "photographic memory." Two of the subjects of this study—Freud (31), Schweitzer (45)—state that their powers of memory began to decline in their twenties; but in his work as a translator Freud continued to read a paragraph, close the book, and write out his translation (31). Macaulay as he grew older needed more effort to learn, but his recall seems to have remained unimpaired (46). Some extraordinary feats of memory are reported of people in their sixties and seventies—Churchill (11), Davis (5), Hadley (interviews), Yardley (52). Nevertheless, failing senses must eventually take their toll—Franklin (21), Lovett (interview), Wells (50).

## H. HEREDITY

Galton describes several families in which members of different generations and collateral relatives had extraordinary memory (23). The materials of the present authors suggest that memory is influenced by heredity, but is not a dominant trait. One person interviewed had excellent visual and auditory learning and recall. She described her father's memory as "photographic," citing his perfect mark on a state examination for a pharmacist's license 20

years after leaving the pharmaceutical school. Her sister's memory, she said, was only fair, and her son's good but not extraordinary. In general the people interviewed reported having parents and children with high general intelligence, but memory varying from good to excellent. A similar pattern covering three or four or more generations is evident in the biographies—Curie (17), Hadley (interviews), Lovett (interview), Macaulay (46). Since memory depends on the acuity of several senses and on different types of mental ability, it evidently reflects the inheritance of both parents.

### I. MEMORY AND ACHIEVEMENT

Memory is best thought of not as a single mental operation but as a series of actions: namely, learning, holding in mind, recall, and recognition [*cf.* Hayes (27)]. Quick learning facilitates the acquisition of information and puts the worker in a position to solve a problem with less effort and more quickly than a less favored individual can. The power to hold information in mind lets the creative person manipulate materials and try out permutations and combinations without the need of concrete models. This is the process of reverie, notably practiced by inventors—Edison (32), Kettering (2), the Wrights (36)—by scientists—Barbour (49), Curie (17), Darwin (19), Michelson (30)—and authors—Maugham (37)—but equally useful to statesmen, lawyers, and executives—Lincoln (43)—all of whom must solve problems in human relations. Faced with a specific detail, the person with a good memory can recall pertinent associations from his prior experiences and from the literature, without the need for laborious checking and cross-checking of indices.

### J. IMPLICATIONS FOR EDUCATION

The relation of superior memory to high achievement has some important implications for education (18). Teachers need to be aware of how easily students with exceptional memory excel on examinations and get good marks—and too often escape intellectual challenge and the satisfaction of hard mental work. Some screening tests, at least of the informal type described here, are in order. There is a possibility that assessment of memory can locate superior intelligence in a disadvantaged child. The early discovery of individuals with relatively great strength of memory is important in guidance, in singling out those who have a potential for high achievement and in encouraging them to prepare to live up to their promise. Finally, all students should be made aware of the role of memory not only in scholastic success but in practical work. They should be taught the old rules—that success depends on having a goal,

mastery of the fundamentals, and persistent effort to achieve. Intention, attention, effort, and persistence are the requisites of learning and of life.

## REFERENCES

1. ALLPORT, G. W. The eidetic image and the after-image. *Amer. J. Psychol.*, 1928, 40, 418-425.
2. BOYD, T. A. Professional Amateur: The Biography of Charles Franklin Kettering. New York: Dutton, 1957. P. 33.
3. BUCHAN, J. Pilgrim's Way. Cambridge, Mass.: Houghton Mifflin, 1940. P. 7.
4. ———. The Thirty-Nine Steps. Boston, Mass.: Houghton Mifflin, 1915. P. 191.
5. BURLINGAME, R. Don't Let Them Scare You: The Life and Times of Elmer Davis. New York: Lippincott, 1961. Pp. 25, 157.
6. ———. Henry Ford. New York: Signet—New American Library, 1956. Pp. 18, 23.
7. BURNS, J. MACG. John Kennedy: A Political Profile. New York: Harcourt, Brace, 1960. P. 262.
8. ———. Roosevelt: The Lion and the Fox. New York: Harcourt, Brace, & World, 1956. Pp. 155, 330.
9. CARPOZI, G. JR. Clark Gable. New York: Pyramid Books, 1961. P. 21.
10. CHOTZINOFF, S. Toscanini: An Intimate Portrait. New York: Knopf, 1956. Pp. 23, 140.
11. CHURCHILL, R. S. Winston S. Churchill. Vol. I: Youth, 1874-1900. Boston, Mass.: Houghton Mifflin, 1966. P. 107.
12. CHURCHILL, W. S. A Roving Commission: My Early Life. New York: Scribner's, 1930. Pp. 25-27.
13. CLAPESATTLE, H. The Doctors Mayo. New York: Pocket Books, 1956. P. 121.
14. CONSIDINE, B. The Babe Ruth Story. New York: Dutton, 1948. P. 56.
15. CONSIDINE, B., & SLOCUM, B. Dempsey. New York: Avon Book Div., Hearst Corp., 1959. P. 125.
16. COX, S. A Swinger of Birches: A Portrait of Robert Frost. New York: New York Univ. Press, 1957. P. 11.
17. CURIE, E. Madame Curie (trans. by V. Sheean). New York: Pocket Books, 1958. Pp. 10, 25, 73, 157-158, 163.
18. CUTTS, N. E., & MOSELEY, N. Teaching the Bright and Gifted. Englewood Cliffs, N.J.: Prentice-Hall, 1957. Pp. 21-22.
19. DARWIN, C. Charles Darwin's Autobiography. New York: Schuman, 1950. Pp. 15, 67, 71.
20. FELDMAN, M. Eidetic imagery in Ghana: A cross-cultural will-o'-the-wisp? *J. of Psychol.*, 1968, 69(2), 259-269.
21. FRANKLIN, B. The Autobiography of Benjamin Franklin (L. W. Larrabee et al., Eds.) New Haven, Conn.: Yale Univ. Press, 1964. P. 156.
22. FULLER, A. A Foot in the Door. New York: McGraw-Hill, 1960. P. 159.
23. GALTON, F. Hereditary Genius. New York: Horizon Press, 1962. P. 20.
24. GREGORY, R. L. Eye and Brain—The Psychology of Seeing. New York: McGraw-Hill, 1966. P. 8.
25. HADLEY, M. Arthur Twining Hadley. New Haven, Conn.: Yale Univ. Press, 1948. Pp. 3, 57, 186.



26. HATCH, A. *The De Gaulle Nobody Knows*. New York: Hawthorn Books, 1960. Pp. 28-29, 35.
27. HAYES, J. R. Memory, goals, and problem solving. In B. Kleinmuntz (Ed.), *Problem Solving—Research, Method, and Theory*. New York: Wiley, 1966. Pp. 149-170.
28. HILTON, C. *Be My Guest*. New York: Popular Library, 1959. P. 63.
29. HUNTER, I. M. L. *Memory* (rev. ed.). New York: Penguin Books, 1964.
30. JAFFE, B. *Michelson and the Speed of Light*. Garden City, N.Y.: Anchor Books, Doubleday, 1960. Pp. 59, 131.
31. JONES, E. *The Life and Works of Sigmund Freud*. New York: Basic Books, 1953. Pp. 55, 57.
32. JOSEPHSON, M. *Edison*. New York: McGraw-Hill, 1959. Pp. 33, 91, 132, 156-158, 442.
33. KELLER, H. *The Story of My Life*. New York: Grosset & Dunlap, 1903. P. 293.
34. KIMBLE, D. P., Ed. *Learning, Remembering, and Forgetting*. Vol. I: *The Anatomy of Memory*. Palo Alto, Calif.: Science and Behavior Books, 1965.
35. LYONS, E. *Our Unknown Ex-President: A Portrait of Herbert Hoover*. New York: Doubleday, 1948. Pp. 84, 95.
36. MCMAHON, J. R. *The Wright Brothers: Fathers of Flight*. Boston, Mass.: Little, Brown, 1930. Pp. 40, 65, 69, 190.
37. MAUGHAM, W. S. *The Summing Up*. New York: Literary Guild of America, 1938. Pp. 1, 82.
38. MILES, C. M. C. *The Early Mental Traits of Three Hundred Geniuses: Genetic Studies of Genius* (Vol. II). Stanford, Calif.: Stanford Univ. Press, 1926. Table 49, p. 2; pp. 323-324, 348.
39. MOOR, P. Maazel. *N.Y. Times Magazine*, Sept. 30, 1962, p. 50.
40. QUINTILIAN. *The Institutio Oratoria: With an English Translation by H. E. Butler* (Vol. 4, Loeb Classical Library). New York: Putnam's, 1920-1922. Pp. 242-243 (Book XI.ii.51).
41. ROOSEVELT, E. *This Is My Story*. Garden City, N.Y.: Dolphin Books, 1937. Pp. 142, 236.
42. RUCH, F. L. *Psychology and Life* (5th ed.). Chicago, Ill.: Scott, Foresman, 1958. P. 350.
43. SANDBURG, C. *Abraham Lincoln* (Vol. 1). New York: Dell, 1959. Pp. 193-194.
44. SCHORER, M. *Sinclair Lewis: An American Life*. New York: McGraw-Hill, 1961. P. 285.
45. SCHWEITZER, A. *Out of My Life and Thought: An Autobiography* (trans. C. T. Campion). New York: Holt, 1933. P. 126.
46. TREVELYAN, G. O. *The Life and Letters of Lord Macaulay* (Vols. 1 & 2). New York: Harper, 1876. Pp. 22-25, 60, 61, 64; 68, 197, 374.
47. WAKSMAN, S. A. *My Life with the Microbes*. New York: Simon & Schuster, 1954. Pp. 48, 49.
48. WECHSLER, D. *The Measurement and Appraisal of Adult Intelligence* (4th ed.). Baltimore, Md.: Williams & Wilkins, 1958. P. 125.
49. WEEKS, E. In *Friendly Candor*. Boston, Mass.: Little, Brown, 1959. P. 150.
50. WELLS, H. G. *Experiment in Autobiography: Discoveries and Conclusions of a Very Ordinary Brain* (Since 1866). New York: Macmillan, 1934. Pp. 13, 14.
51. *WHO'S WHO IN AMERICA*. Chicago, Ill.: Marquis, 1967. P. 1299.

52. YARDLEY, H. O. The Education of a Poker Player. New York: Simon & Schuster, 1957. Pp. 88-89.

*2 Fifth Avenue  
Hotchkiss Grove  
Branford, Connecticut 06405*

*Rural Route 1, Box 60  
Sandy Hook, Connecticut 06482*





## ABASEMENT SCORES AND ADJUSTMENT OF NEUROPSYCHIATRIC PATIENTS\*

*Texas Research Institute of Mental Sciences*

---

ALLEN GOSS<sup>1</sup>

---

### A. INTRODUCTION

Increased interest between the need scores of the Edwards Personal Preference Schedule (EPPS) and other aspects of total life adjustment (4) is being reflected in the literature. Rosenkrantz and O'Halloran (7) have recently explored the relationship between the abasement score of the EPPS and adjustment in a college population. Theoretically, high self-abasement is expected to accompany maladjustment and low self-abasement is associated with health or good adjustment. When 26 psychologists ranked the 15 need variables of the EPPS with regard to their adjustive consequences (7), abasement was on the bottom. Goodstein and Heilbrun (2) had college undergraduates rate these same 15 variables on personal desirability and social desirability; and ranks of 14 and 11 respectively were given to abasement. Izard (5) has found a significant negative correlation between abasement score and academic performance for males, but not for females. There is, however (6), a significant decrease in abasement scores over the four years of college for both sexes.

It would seem that if a relation between self-abasement and maladjustment exists and if the EPPS abasement score measures an important aspect of the concept of self-abasement, this relation would be most easily observed in a neuropsychiatric population. Furthermore, if a neuropsychiatric population was divided so that those who benefited from a rehabilitation program could be separated from those who did not, abasement scores should be higher for the nonbenefit group.

### B. METHOD

The investigation was based on the records of 58 male neuropsychiatric patients accepted to a Vocational Rehabilitation Program at the VA Hospital, Houston, Texas. The population comprised 18 alcoholics, 20 anxiety or depres-

---

\* Received in the Editorial Office, Provincetown, Massachusetts, on September 9, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> Data for this study was collected while the author was a psychology trainee and postdoctoral fellow at the VA Hospital in Houston.

sion reactions, nine with physical disabilities, and 11 schizophrenics. The mean age of the population was 41 and the populations did not differ significantly with respect to 10 demographic variables (3). Of this total population 33 gained employment and 25 were discharged from the ward without employment or were transferred back to traditional treatment oriented wards.

### C. RESULTS AND DISCUSSION

The results show that there is little difference between the percentile scores of neuropsychiatric patients and the General Adult Group norms (1). An analysis of variance between the success and fail populations indicates that these populations do not differ with respect to their abasement scores. In fact, in three of four cases the population which gained employment had higher abasement scores. Other EPPS score dimensions were useful in differentiating the success-fail population and provided a model capable of predicting success well above the base rate (3).

TABLE 1

PERCENTILE SCORES OF DIAGNOSTIC POPULATIONS IN TERMS OF REHABILITATION OUTCOME

Group	Success	Fail	F.	P.
Total	51.30 <sup>a</sup>	47.88	.18	.67
Alcoholic	50.50	30.37	1.83	.20
Anxiety-depression	48.85	59.43	.56	.46
Physical disability	39.75	36.00	.04	.84
Schizophrenia	57.67	55.50	.01	.92

<sup>a</sup> Percentile Scores from the General Adult Group norms.

These results are quite similar to the results of Rosenkrantz and O'Halloran (7), since their college population was not differentiated on adequacy of adjustment by the abasement scores. Their results, like those in the present study, indicated a tendency for the "better" adjusted group to score higher on the abasement index instead of lower, as traditional clinical reasoning seems to suggest. Two explanations for this difference are (a) the abasement scale of the EPPS does not measure critical aspects of the concept of self-abasement, and (b) self-abasement is not an important aspect of adequate vocational rehabilitation. At present there exists a need for empirically derived information and concepts related to criteria that are behaviorally defined.

### D. SUMMARY

High self-abasement is often expected to accompany maladjustment, and low self-abasement is associated with health or a good life adjustment. This study investigated the relation between the abasement score of the EPPS and adjust-

ment of patients on a vocational rehabilitation ward. The results showed no significant difference on the abasement score of the EPPS between patients who got jobs and patients who did not get jobs or between patients and the general adult norms. The abasement scale of the EPPS does not measure important aspects of the concept of self-abasement and/or self-abasement is not critical for adequate life adjustment.

## REFERENCES

1. EDWARDS, A. L. Manual for the Edwards Personal Preference Schedule. New York: Psychological Corp., 1959.
2. GOODSTEIN, L. D., & HEILBRUN, A. B. The relationship between personal and social desirability values of the Edwards Personal Preference Schedule. *J. Consult. Psychol.*, 1959, 23, 183.
3. GOSS, A. Predicting work success for patients on an industrial rehabilitation ward in a neuropsychiatric setting. Doctoral dissertation, University of Texas; available, University of Michigan Microfilms, No. 66-14381, Ann Arbor, Michigan, 1966.
4. HEILBRUN, A. B. Social-learning theory, social desirability, and the MMPI. *Psychol. Bull.*, 1964, 61, 377-388.
5. IZARD, C. E. Personality Characteristics (EPPS), level of expectation, and performance. *J. Consult. Psychol.*, 1962, 26, 394.
6. ———. Personality change during college years. *J. Consult. Psychol.*, 1962, 26, 482.
7. ROSENKRANTZ, P. S., & O'HALLORAN, W. S. Are abasement scores predictive of adjustment? *J. of Psychol.*, 1965, 61, 33-37.

*Texas Research Institute of Mental Sciences*  
1300 Moursund Avenue  
Houston, Texas 77025





# PSYCHODYNAMIC FACTORS IN THE EXPERIMENTAL INVESTIGATION OF HYPNOTICALLY INDUCED EMOTIONS WITH PARTICULAR REFERENCE TO BLOOD GLUCOSE MEASUREMENTS\*<sup>1</sup>

*Institute for Research in Hypnosis, New York, New York*

MILTON V. KLINE<sup>2</sup> AND MORTON LINDER<sup>3</sup>

The use of hypnosis to induce a variety of emotional responses has been reported by many investigators in connection with specific studies of behavioral responses to hypnosis. Such studies have assessed the validity of induced emotional responses as well as the characteristic psychophysiological reactions produced (4). The present investigators have been involved in studies with diabetic and normal subjects in whom hypnotically induced emotions have been studied in relation to blood glucose levels. In our study, as well as others, it has been found that hypnotically induced emotional stress usually fails to elevate blood glucose levels significantly and consistently (12).

It was postulated that one of the possible explanations for the failure of hypnotically induced stress to produce an elevation in blood glucose levels was that physiologically active homeostatic mechanisms in controlling blood glucose appear to be unaltered in both normal and mild diabetic subjects despite apparent intense hypnotically induced stress.

Vandenbergh, Sussman, and Titus (11) found it difficult to explain a slight decrease in blood glucose in their studies on hypnotically induced stress. In view of the current findings, it is important to be able to account for what has been observed and how it relates both to the nature of the experimental investigations and to the psychophysiological and psychodynamic constructs which may be involved.

An evaluation of several studies dealing with blood sampling indicates that venal puncture in itself frequently produces more anxiety of a disrupting nature than does hypnotically induced stress designed to be intense in nature. In

\* Received in the Editorial Office, Provincetown, Massachusetts, on September 10, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This paper represents part of a research investigation presented at the 7th Annual International Psychosomatic Congress, Rome, Sept. 11-16, 1967.

<sup>2</sup> Executive Director, The Morton Prince Clinic for Hypnotherapy, The Institute for Research in Hypnosis, New York City.

<sup>3</sup> Research Consultant in Medicine, The Morton Prince Clinic for Hypnotherapy, The Institute for Research in Hypnosis, New York City.

the Vandenberg, Sussman, Titus study (11), it is interesting to note that they state that "each patient reported anxiety on the first control day, due to an initial exposure to the intravenous drip in his arm."

It seems that, although experimental investigators are attempting to induce acute emotional stress through the use of hypnosis, in effect they may inadvertently be achieving with certain subjects a therapeutic-like response and a discharge of affect which produces first an acute excitation followed rapidly by a state of internalized homeostatic equilibrium.

The technique of abreaction which has been utilized is described most clearly by Vandenberg *et al.* as the hypnotic method for the induction of emotional stress. Abreaction as a term originates in the theatre (Aristotle) (1) and permits the patient to act out and revivify, as well as discharge, meaningful memory and correlated behavioral experiences. Abreaction originated as a therapeutic device alone and, while it has been insufficient very often to produce lasting effects in terms of the treatment of neurotic or psychotic disturbances, it, nevertheless, has had a significant role in the history of psychotherapy and continues to be of importance.

Shorvon and Sargent (10) describe an abreaction as a "process of reviving the memory of a repressed, unpleasant experience and expressing, in speech and action, the emotions related to it, thereby relieving the personality of its importance." The emphasis in this definition is clearly the *release of the emotions*.

Clinicians and therapists who have had considerable experience with abreactions are familiar with the fact that the discharge of released emotions is usually followed by a state of exhaustion in which the patient shows a loss of muscle tone and is quiet. The quietness produces a degree of relaxation which is frequently so great as to have not only momentary but lasting therapeutic value for a varying period of time.

Sargent (9) compares this state of exhaustion to the transmarginal inhibition of Pavlov in which dogs following an emotional trauma lose the previously acquired conditioned reflexes or show a complete reversal of conditioned reaction.

Gellhorn and Looftbourrow (2) suggest that the physiologic processes underlying the excitatory abreaction are an intensive, hypothalamic-cortical discharge followed by a period of lessened cortico-hypothalamic-cortical relations. The severe hypothalamic-cortical discharge due to extreme excitation is induced by encouraging the patient to relive the traumatic experience. The subsequent reduction of the hypothalamic-cortical discharge is attributed to excessive adrenomedullary secretion which has been shown to depress the



sympathetic division of the hypothalamus, and possibly also to an exhaustion of the hyperreactive division of the hypothalamus. The therapeutic usefulness of the abreaction would appear to depend on whether maximal excitation followed by exhaustion is attained. Clinical experience with continuous abreactive experiences has been demonstrated to relieve markedly a number of pathological symptoms unresponsive to a single abreaction (4).

Hypnosis involves the activation of strong transference phenomena between hypnotist and patient or subject, and even within the confines of investigational procedures the therapeutic meaningfulness of the hypnotic relationship has been repeatedly observed and reported (4, 5). Thus, within the "transference-relaxation" aspect of the hypnotic process, there is induced abreaction which can produce "therapeutic value" due to its discharge-exhaustion syndrome and, at the same time, involve the subjects in a process of desensitization. They are abreacting painful material within the confines of a very favorable and accepting relationship. This certainly fits into the framework of reciprocal inhibition as described by Wolpe and behavioral therapists (14).

It is, thus, not surprising that blood glucose levels are frequently not affected or, in fact, may be somewhat depressed by the use of investigational techniques which are heavily influenced by a hypnotic transference, the presence of therapeutic exhaustion, and the autonomic nervous system discharge which follows within a behavioral process involving desensitization and the effects of reciprocal inhibition.

Insulin is unique in its ability to lower blood glucose levels. All other known hormones that can affect blood glucose levels do so in one direction, elevating them, and oppose the action of insulin in so doing. The major hormones involved are hydrocortisone-cortisol, A.C.T.H., growth hormone, and epinephrine; and interestingly enough, their levels are all increased with anxiety. The actions of glucagon are more complex, since it elevates blood glucose levels and increases insulin release simultaneously. Nevertheless, its primary effect upon blood glucose levels is to elevate them. Regardless of the mechanism by which blood glucose levels are changed, be it via production or utilization variations, it seems clear that increased blood sugar levels are associated with increased cortisol (and epinephrine) level, and *vice versa*. One can, therefore, deduce whether a situation is anxiety producing or anxiety relieving by the changes in hormone or blood glucose levels.

Many investigators have occupied themselves with this matter. Levitt, Persky, and Brady (6) found that plasma cortisol levels seemed to go up with tension and down with relief of tension. This included elevation with an intravenous placebo at initial injection, and little elevation or, indeed, reduction

with further intravenous placebo injections. Persky *et al.* (8) showed that the character of the stimulus exerts a significant effect upon plasma cortisol levels and that the direction of change in the hormone level depended in part upon its initial level, so that high initial levels (reflecting initial anxiety) often result in lower subsequent levels, regardless of the experimental design. This was first shown by Wilder (13), who stated that there is a negative and highly significant correlation between the initial hormone level and the change in level, and he called this phenomenon his *Law of Initial Values*. Persky (7) noted that in order to avoid the countervailing influence of Wilder's Law, it is necessary to employ a sample population with a low initial cortisol level with small variability consistently.

The results of Levitt, Persky, and Brady (6) and of other investigators indicate that an elevation in anxiety results in an elevation in cortisol level, while an elevation in cortisol level (endogenous or exogenous) brings about an increase in anxiety-proneness. It is possible that these variables are related in some servo-regulatory system. These workers found that their anxiety producing hypnotic procedures resulted in consistent elevation in anxiety levels as revealed by clinical ratings, psychological test indices, and by elevation in cortisol and A.C.T.H. levels. They state that, in general, the magnitude of cortisol elevation in response to psychological stress is proportional to the intensity of the stress. Handlon (3) also observed that anxiety reduction leads to a diminution in plasma cortisol level.

The procedure utilized by Vandenberg, Sussman, and Titus (11) was described as abreacting, and one can readily explain the reduced blood glucose levels on the physiological grounds previously discussed: namely, a reduction in anxiety from the abreaction, with a consequent reduction in hormone levels causing the reduced blood glucose levels. In addition, the factor of hydration with 1500 cc of water and intravenous infusions must be taken into account.

These observations do not negate the meaningfulness of emotions produced through hypnosis. Rather they emphasize the psychophysiologic avenues through which such emotions are channeled and the effect they may have on psychosomatic functioning and those attempts to measure the by-product of such an emotional process. Polygraphic monitoring of hypnotically induced emotions, particularly upper thoracic and abdominal respiratory patterns and GSR levels, also reflect the immediacy of the emotional response.<sup>4, 5</sup> This, however, is extremely short lived and despite the "disturb-

<sup>4</sup> Kline, M. V.: (unpublished) paper presented to the Departments of Psychiatry and Psychology, St. Vincent's Medical Center, New York City, April, 1967.

<sup>5</sup> Kline, M. V., Research program, presented to the Graduate Faculty in Psychology, St. John's University, New York City, May, 1967.

ing" aspects of it frequently leads even experimental subjects, as well as patients, to report that they "feel better" after the experience. Here again, one observes the type of relaxation associated with abreactive experience.

It is clear that not all emotional expression involves the same psychophysiological mechanisms, and those that occur within the framework of the hypnotic transference and the hypnotic state may constitute rather significant differences from those emotions that occur under naturally occurring life involvements and life experiences. This contrast and discrepancy is of importance in viewing the nature of emotionally induced behavior and, particularly, those therapeutic approaches which may be most valuable in dealing with them.

## REFERENCES

1. FENICHEL, O. On acting. *Psychiat. Quart.*, 1946, 15, 144-160.
2. GELLHORN, E., & LOOFBOURROW, G. N. Emotions and Emotional Disorders. New York: Harper & Row (Hoeber Medical Division), 1963.
3. HANDLON, J. H. Hormonal activity and individual responses to stresses and circumstances in everyday living. In *Physiological Correlates of Psychological Disorder*, Roessler & Greenfield, Eds. Madison: Univ. Wisconsin Press, 1962.
4. KLINE, M. V., Ed. *Clinical Correlations of Experimental Hypnosis*. Springfield, Ill.: Charles C Thomas, 1963.
5. ———. *Psychodynamics and Hypnosis*. Springfield, Ill.: Charles C Thomas, 1967.
6. LEVITT, E. E., PERSKY, H., & BRADY, J. P. Hypnotic Induction of Anxiety. Springfield, Ill.: Charles C Thomas, 1964.
7. PERSKY, H., GRINKER, R. R., et al. Adrenal cortical function in anxious human subjects: Plasma level and urinary excretion of hydrocortisone. *Arch. Neurol. & Psych.*, 1956, 76, 549-558.
8. PERSKY, H., HAMBURG, D. A., et al. Relation of emotional responses and changes in plasma hydrocortisone level after stressful interview. *Arch. Neurol. & Psych.*, 1958, 79, 434-447.
9. SARGENT, W. *Battle for the Mind*. London, England: Heinemann, 1957.
10. SHORVON, H. J., & SARGENT, W. Excitatory abreaction. *J. Ment. Sci.*, 1947, 93, 709-732.
11. VANDENBERGH, R. L., SUSSMAN, K. E., & TITUS, C. Effects of hypnotically induced acute stress on carbohydrate and liquid metabolism in patients with diabetes mellitus. *Psychosomat. Med.*, 1966, 4(1), 382-390.
12. WELLER, C., LINDER, M., NULAND, W., & KLINE, M. V. The effects of hypnotically induced emotions on continuous, uninterrupted blood glucose measurements. *Psychosomat. Med.*, 1961, 5(1), 38.
13. WILDER, J. Zft. fur gesamte. Neurol. & Psychol., 1931, 137-338, as quoted in 1.
14. WOLPE, J. *Psychotherapy by Reciprocal Inhibition*. Stanford, Calif.: Stanford Univ. Press, 1958.

345 West 58th Street  
New York, New York 10019





## THE USE OF FIGURE DRAWINGS TO ASSESS RELIGIOUS VALUES\*<sup>1</sup>

*Department of Psychology, University of Tennessee*

C. CLIFFORD ATTKISSON, LEONARD HANDLER, AND RAYMOND R. SHRADER

### A. INTRODUCTION

In a recent article Dennis (1) concluded that the Draw-A-Man Test is a valid instrument in reflecting religious values. Dennis theorized that if persons who dedicate their lives to piety and religious service can be considered to possess strong religious values, it would follow that these highly religious persons should include religious content in their drawings. Dennis collected drawings from 100 nuns (predominantly Roman Catholic) in Lebanon and found that 74 per cent of the nuns included religious symbols in their drawings.

While Dennis' data seem rather clear-cut, the present authors argue that his interpretations must be qualified, for two reasons. First, Dennis' sample of nuns represents a highly selected and socially secluded group. Second, there is no evidence that these findings can be generalized to the many nonprofessional religious persons who also hold religious values as central. The present authors felt that a better test of Dennis' hypothesis concerning the ability of the DAP to reflect religious values would require that drawings be collected from highly religious adults who, nevertheless, are integrated into the mainstream of the secular world.

The purpose of this study, therefore, was to determine whether individuals with varying degrees of involvement in religious life represent this value in their drawings, as Dennis suggests. If the drawings show concern with religious people, symbols, activities, and places, then the power of the Draw-A-Person technique to demonstrate religious values would be supported more substantially than in the Dennis article.

---

\* Received in the Editorial Office, Provincetown, Massachusetts, on September 11, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> The writers would like to thank the following people who were kind enough to assist in the data collection: Sister Mary Raphael Lozes, Drs. John M. Gessell and J. Lyn Elder, Hubert T. Davis, Grayson L. Tucker, John M. Lowe, Frank Erikson. They would also like to thank Steve Saywell who assisted in various phases of the study.

## B. METHOD

Drawings were collected from Baptist, Episcopalian, Presbyterian, and Church of Christ theology students; Methodist undergraduate students who comprised the board of directors of a religious organization; Roman Catholic priests and theology students; and Roman Catholic Dominican nuns ( $N = 215$ ). The Ss' personal commitment, their choice of profession or choice of religious order, and their pursuit of religious values all reflected a high degree of religious involvement.

Table 1 indicates the number of Ss in each group, the location of the data collection, the age range and mean age of each group, and the number of males and females in each group. The mean age of the total sample was 25 years (range = 18-50), and the subjects were predominantly male (158 males *vs.* 57 females).

TABLE 1  
RELIGIOUS SUBGROUPS COMPRISING THE TOTAL SAMPLE

Denomination and location	<i>N</i>	Age range	Mean age	Male	Female
Episcopal (Tennessee)	29	—	28	29	0
Presbyterian (Kentucky)	45	23-50	27	45	0
Baptist (California)	22	24-49	30	22	0
Church of Christ (Tennessee)	62	18-50	21	38	24
Methodist (Virginia)	14	18-26	21	9	5
Roman Catholic priests (Tennessee)	15	18-26	22	15	0
Roman Catholic nuns (Louisiana)	28	21-48	29	0	28
Total	215	18-50	25	158	57

The drawings were group administered and were collected either by the authors or by fully instructed supervisory personnel from within the groups. Care was taken to avoid any mention of the purpose of the drawing request.

Ss were supplied with two sheets of  $8\frac{1}{2} \times 11$  unlined paper and a number two pencil. *E* then stated: "On the first sheet of paper I would like you to draw a person. Please be sure that you draw a whole person and not just a head. We are not interested in how good an artist you are. This is not a test



of artistic ability." Whenever *E* was questioned, he responded with the following statement: "Draw the figure any way you like. Just be sure to draw the whole person and not just the head."

When the first drawing was completed it was turned over, and *S* was then instructed: "On the second sheet of paper, draw a person which is opposite in sex from the figure you drew on the first sheet of paper. Be sure to draw a whole person."

When the two drawings were completed *E* stated: "We want to be able to pair the two drawings you made. Therefore, pick out some identifying mark or symbol and put it on both of your drawings."

### C. RESULTS

A total of 430 drawings was collected. Each drawing was examined for the presence of religious content by the first two *Es*. Prior to the data collection it was decided that any indication of a religious symbol, person, activity, or place would be scored as religious content. The two *Es* obtained perfect agreement. To guard against rating bias by the *Es*, a third rater, one who had no knowledge of the hypotheses, also rated the drawings. The per cent agreement between the third rater and each of the first two raters was 98 per cent.

Table 2 presents the percentage of *Ss* whose drawings did and did not include religious content. Only seven of the 215 *Ss* (three per cent) made

TABLE 2  
RELIGIOUS CONTENT IN THE DRAW-A-PERSON TEST FOR THE RELIGIOUS SUBGROUPS

Denomination	Religious content	Nonreligious content
	0	29
Episcopal	0	45
Presbyterian	0	22
Baptist	4	58
Church of Christ	0	14
Methodist	1	14
Roman Catholic priests	2	26
Roman Catholic nuns	7	208
Total		

drawings containing religious content in one of their two drawings. Only one of these seven *Ss* included religious content in both drawings. Thus there were eight drawings with religious content out of a total of 430 drawings (less than two per cent). Nine per cent of the *Ss* chose religious symbols to code their protocols.

In every case the drawing containing religious content was the first one made by the *S*. The actual religious content included semicomical doomsday

figures, figures of Christ, contemporary persons holding bibles, a nun, and one drawing which was highly elaborated with symbolic religious comments. Four drawings containing religious content were done by male Ss and three were done by females.

A wide range of nonreligious content was represented: e.g., men in ordinary business suits and in collegiate dress, women in ordinary dresses and typical collegiate dress, nudes (present in all but the Roman Catholic groups), and drawings of beatniks. Other content represented included farmers, sailors, athletes of many kinds (including golfers, football, baseball, and basketball players), superman and superwoman, soldiers (including several figures firing guns and otherwise fighting), women in bathing suits, elaborately dressed women, women in low cut dresses and "well-endowed" women, spacemen, cowboys and cowgirls, etc.

There was a wide variety of content represented in the nuns' drawings, including a goodly proportion of beatniks (present most often in the nuns' drawings compared with the other groups), Indians, cheerleaders, athletes, a figure clad in a bathing suit, and a figure in a military uniform.

#### D. DISCUSSION

The most interesting finding is, of course, that 97 per cent of the highly religious individuals in this sample did not produce drawings containing overt religious content. Furthermore, as indicated in Table 2, no drawings manifesting religious content were produced by representatives of the larger Protestant groups: Methodists, Baptists, Presbyterians, and Episcopalians. All drawings having religious content were obtained from the remaining three groups: The Church of Christ (four drawings), Roman Catholic priests and theology students (one drawing), and Roman Catholic nuns (two drawings).

This study made no attempt to sample all major religious groups, nor does it propose to generalize its findings to all such groups. Nevertheless, it is interesting to note that the majority of religious drawings were obtained from fundamental and conservative groups. None of the Ss in this study were as socially and geographically isolated as those in Dennis' study. Therefore, the present writers feel that the factors of environmental detachment, social seclusion, and perhaps a highly authoritarian social hierarchy are probably important variables in determining the production of religious content in drawings of the human figure.

The data cast doubt on Dennis' (1, 2) contention that the Draw-A-Man Test is a valid instrument to reflect religious values, at least for those living in the USA. Rather than conclude, as did Dennis, that "... few eleven-

twelve-, and thirteen-year-olds show concern with religious men, religious places, religious activities" (2, p. 134), the present authors conclude that the Draw-A-Man Test is not a useful instrument to tap religious values in most groups.

The fact that Dennis (1) obtained religious drawings from 74 per cent of the nuns used in his sample probably indicates that the test does tap religious values in some *extreme* groups. However, for all *practical* purposes the test seems of little value in tapping general religious values.

While the presence of religious symbols in a drawing may indicate that the artist has religious values, the absence of religious symbolism does not necessarily indicate the absence of these values. In a similar manner, since the presence of religious symbolism in drawings is such a rare occurrence, a clinician can feel relatively safe in hypothesizing that for this person religion is probably a central value.

#### E. SUMMARY

The present study was a test of Dennis' hypothesis concerning the ability of the DAP to reflect religious values. Drawings were collected from Baptist, Episcopalian, Presbyterian and Church of Christ theology students; Methodist undergraduate students who comprised the board of directors of a religious organization; Roman Catholic priests and theology students; and Roman Catholic Dominican nuns ( $N = 215$ ). Only seven of the 215 Ss (three per cent) made drawings containing religious content. These results cast doubt on Dennis' contention that the DAP is a valid instrument to reflect religious values, at least for those living in the USA.

#### REFERENCES

1. DENNIS, W. *Group Values Through Children's Drawings*. New York: Wiley, 1966.
2. DENNIS, W., & URAS, A. The religious content of human figure drawings made by nuns. *J. of Psychol.*, 1965, 61, 263-266.

*The Psychological Clinic*  
*The University of Tennessee*  
*1303 West Cumberland Avenue*  
*Knoxville, Tennessee 37916*





## SEMANTIC SPACE IS (APPROXIMATELY) BIPOLAR\*

*Department of Psychology, University of California, Los Angeles*

---

P. M. BENTLER

---

### A. INTRODUCTION

One of the most useful measurement techniques for human experimental, social, and applied areas of psychology has been the semantic differential procedure of Osgood, Suci, and Tannenbaum (7). As is well known, these authors found that the connotative meaning of objects and concepts could be summarized adequately by three major dimensions: evaluation, activity, and potency. Each of these three dimensions has been defined by several seven-point bipolar semantic differential rating scales. For example, the evaluative dimension is represented by such scales as good-bad, pleasant-unpleasant, and nice-awful. The activity dimension is represented by such scales as active-passive, fast-slow, and agitated-calm. The potency dimension is depicted by such scales as strong-weak, heavy-light, and hard-soft.

A crucial assumption used by Osgood and his colleagues in arriving at his three-dimensional description of meaning was that semantic space is bipolar: that is, that each dimension is represented by polar oppositional terms, usually antonyms. Thus, rating scales used to measure the dimensions characteristically are defined at opposite ends of each scale by polar oppositional terms, such as the ones mentioned above. While Carroll (2) criticized Osgood *et al.* for the often arbitrary definition of opposites, it was not until a short time ago that the question of the bipolarity of semantic space was investigated experimentally. Green and Goldfried (3) argued that Osgood *et al.* built bipolarity of the factor structure of semantic space into their procedures when presenting subjects with polar-oppositional rating scales for the description of concepts. It appeared quite possible that the bipolarity found for semantic space represented primarily the effects of experimental procedures rather than some intrinsic quality of subjects' responses. Since the usual rating scale instructions did not allow subjects to check both ends of a scale—for example, in order to describe a concept as both good *and* bad—Green and Goldfried felt the procedures did not even allow Osgood's implicit assumption to be tested. Fur-

---

\* Recommended by Ernest R. Hilgard of the Editorial Board, and received in the Editorial Office, Provincetown, Massachusetts, on September 17, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.



## SEMANTIC SPACE IS (APPROXIMATELY) BIPOLAR\*

*Department of Psychology, University of California, Los Angeles*

---

P. M. BENTLER

---

### A. INTRODUCTION

One of the most useful measurement techniques for human experimental, social, and applied areas of psychology has been the semantic differential procedure of Osgood, Suci, and Tannenbaum (7). As is well known, these authors found that the connotative meaning of objects and concepts could be summarized adequately by three major dimensions: evaluation, activity, and potency. Each of these three dimensions has been defined by several seven-point bipolar semantic differential rating scales. For example, the evaluative dimension is represented by such scales as good-bad, pleasant-unpleasant, and nice-awful. The activity dimension is represented by such scales as active-passive, fast-slow, and agitated-calm. The potency dimension is depicted by such scales as strong-weak, heavy-light, and hard-soft.

A crucial assumption used by Osgood and his colleagues in arriving at his three-dimensional description of meaning was that semantic space is bipolar: that is, that each dimension is represented by polar oppositional terms, usually antonyms. Thus, rating scales used to measure the dimensions characteristically are defined at opposite ends of each scale by polar oppositional terms, such as the ones mentioned above. While Carroll (2) criticized Osgood *et al.* for the often arbitrary definition of opposites, it was not until a short time ago that the question of the bipolarity of semantic space was investigated experimentally. Green and Goldfried (3) argued that Osgood *et al.* built bipolarity of the factor structure of semantic space into their procedures when presenting subjects with polar-oppositional rating scales for the description of concepts. It appeared quite possible that the bipolarity found for semantic space represented primarily the effects of experimental procedures rather than some intrinsic quality of subjects' responses. Since the usual rating scale instructions did not allow subjects to check both ends of a scale—for example, in order to describe a concept as both good and bad—Green and Goldfried felt the procedures did not even allow Osgood's implicit assumption to be tested. Fur-

---

\* Recommended by Ernest R. Hilgard of the Editorial Board, and received in the Editorial Office, Provincetown, Massachusetts, on September 17, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

thermore, they felt that many adjectival dimensions could not reasonably be considered as bipolar; as an illustration, they described the "loneliness" continuum as a unipolar continuum possessing no reasonable opposite.

The assumption of bipolarity was tested by Green and Goldfried as follows. A single-adjective scale version of the semantic differential was constructed so as to possess the graphic rating scale characteristics of the usual semantic scales. However, instead of labeling the two ends of a scale by oppositional terms, only a single adjective was presented for each scale. This adjective was written slightly above the rating scale, and subjects were instructed to mark the extent to which that particular adjective was related to the concept to be described. This format allowed concepts to be rated as usual, but in addition made it possible to present antonym adjectives on separate rating scales. Thus, subjects could describe concepts by indicating separately the extent to which they believed a concept was described by opposite adjectives. A number of subjects were asked to describe a number of concepts in terms of 60 adjectives. The adjectives were chosen to represent each of the two poles of each of the three semantic factors.

Correlation coefficients described the extent of opposition of polar opposite adjectives. Thus, if "good" was indeed the opposite of "bad," the correlation between the two scales represented by these adjectives, across subjects and concepts, should be highly negative. If polar opposites were not intrinsically opposite, then the correlations between presumed opposite pairs might be close to zero or even positive. A detailed analysis of the correlation matrices derived from their data led Green and Goldfried to the conclusion that semantic space was *not* in general bipolar, although certain adjective pairs did appear to be functional opposites. The implication of these results is profound: describing a concept in terms of Osgood's system represents hardly more than a description using an arbitrary three-dimensional bipolar system. It would appear as if the semantic differential should be abandoned.

Conclusions regarding the uselessness of the bipolar semantic differential from studies based upon a unipolar semantic differential must, however, assume that no special conditions affected the results of the unipolar studies. Unfortunately, the Green and Goldfried studies are open to criticism on the basis of the possible existence of a response style. A recent review by Hamilton (4) has indicated that rating scales in general are quite susceptible to an extremity response style. If an extremity response bias existed in the Green and Goldfried study, its effect would be to attenuate the potentially high negative correlation between polar oppositional terms. This effect would be particularly noticeable if the extremity bias represented primarily an acquiescence



tendency; then subjects might indicate that a descriptive adjective is very positively related to a particular concept, but that its opposite may also be very positively related to the concept. Green and Goldfried provided no evidence regarding the existence or nonexistence of response styles for scales administered according to their format. They do not raise the possibility that polar oppositional semantic tendencies might be negated by the existence of nonoppositional, or one-sided, irrelevant response tendencies. The purpose of this study is to evaluate whether or not semantic scales tend to be bipolar when response styles are controlled.

Should semantic space turn out not to be bipolar, even when response styles are taken into account, then it becomes imperative to provide some kind of a measuring instrument for the assessment of connotative meaning which might replace the traditional semantic differential. While Green and Goldfried attempted to demonstrate that semantic space was not bipolar, they failed to provide an alternative set of unipolar scales to assess various directional vectors in multidimensional space of meaning. An additional purpose of this study is thus to provide for a thorough assessment of the possible unidirectional vectors in semantic space.

## B. METHOD

An adjective version of the semantic differential was constructed. Three hundred fifty-two adjectives were collected from semantic differential literature to assess the three semantic dimensions. The evaluative dimension was represented by 141 adjectives, the activity dimension by 106, and the potency dimension by 105 adjectives. Adjectives represented each of the two poles of each of these three major semantic factors. The entire list of 352 adjectives was given to a subject, who was asked to describe the emotional meaning of a particular concept in terms of the adjectives. Two hundred concepts were selected for the study from the published semantic differential profiles of Jenkins, Russell, and Suci (6) and Heise (5). Concepts were chosen systematically to represent each of the eight possible combinations of the positive and negative poles of evaluation, activity, and potency. One hundred male and 100 female undergraduate sophomores were used as subjects for the study.

The basic data from the study could be represented in a 200 by 352 cell matrix, representing the response of all subjects to all items. Data analysis consisted of three major phases. The first phase consisted of a three-step multivariate analysis of the data matrix just described. The first analysis attempted to determine a cluster of positive evaluation ( $E+$ ) and a cluster of negative evaluation ( $E-$ ) adjectives; it consisted of an analysis of the eval-

uation adjectives alone. The second analysis consisted of an attempt to define positive (A+) and negative (A-) ends of the activity dimension; and a corresponding third analysis consisted of an attempt to define positive (P+) and negative (P-) ends of the potency dimension. Each of these analyses consisted of an application of a scaling procedure called multidimensional homogeneity scaling (1). This technique, which is essentially a nonmetric factor analysis, purports to assess the latent ordinal dimensions associated with the observed responses to test items. Like Guttman's or Loevinger's approach to scaling, it aims at identifying ordinal scales of responses; like Thurstone's multiple factor analysis, it attempts to determine all common latent dimensions. Items are first scored dichotomously, followed by the computation of interitem homogeneities to assess the extent to which pairs of items represent an ordinal scale. These homogeneities are  $\phi/\phi$  max or  $\phi/|\phi|$  min coefficients. A matrix coefficient of homogeneity is then defined, and items are weighted to maximize total test homogeneity, or scalability. The weighting is linear; consequently, successive orthogonal homogeneous dimensions can be identified. In each of the three analyses two major dimensions were identified which were rotated by the varimax procedure.

The second phase of the study consisted of finding adjectives that marked each of the six major dimensions which emerged from the scaling procedure. Twenty-five adjectives were chosen to represent each of the two poles for each of the three multivariate analyses. Twenty-five adjectives thus assessed a given direction for a particular semantic scale. Thus, for example, the adjectives *good, beautiful, pleasant, sincere*, among others, assessed positive evaluation; the adjectives *dangerous, troublesome, unpleasant, cruel*, among others, represented negative evaluation; *active, alive, dynamic, moving, lively*, represented positive activity; *quiet, cold, old, slow, passive*, and others, represented negative activity; *powerful, important, strong, serious, compelling*, represented positive potency; and *smooth, delicate, easy, simple, light*, and others, represented negative potency. Since each pole of each of the semantic scales was now represented by 25 adjectives, it was possible to determine a total score for a particular concept by simply counting up the number of adjectives which were felt by a subject to be descriptive of that concept. Thus, a given concept could be represented by scores ranging from zero to 25 for any given pole on each of the original Osgood dimensions.

The final step in the procedure consisted of calculating the correlations between the total scores assigned to concepts on each of the two poles of the three traditional dimensions. These correlations then yielded a six by six correlation matrix, correlations being calculated across the 200 subjects and concepts being evaluated.

The crucial analysis regarding the bipolarity of semantic space is represented in the correlation matrix just described. If only the specific semantic content determined the total score on a particular scale, then the two scales for a given semantic factor—which presumably measure opposite ends of the same factor—should be highly negatively correlated. Thus, positive (E+) and negative (E-) evaluation should be highly negatively related, as would positive activity (A+) and negative activity (A-), and also positive potency (P+) and negative potency (P-). In addition, cross-factor correlations should be reasonably low—certainly lower in absolute value than the absolute value of a correlation of the polar opposites. In contrast, if the semantic differential is not bipolar or if a response style is a major determinant of scale scores, then polar opposites may be virtually unrelated to each other. Lack of bipolarity may stem from the nonoppositional nature of connotation, or because genuine content response tendencies and response style tendencies work against each other. If only a trait endorsement or acquiescent tendency determines the responses without any effect from the connotative content whatsoever, then all scales would be highly positively related, including polar opposites. This is because the total score for a particular concept would represent simply the number of adjectives checked on each of the six scales and the number checked would be representative of a checking tendency or acquiescent tendency on the part of a subject. If both response style and connotative content determine the responses, the correlations should represent some compromise between these two opposing tendencies.

### C. RESULTS

One of the most important questions to be answered by the data is whether or not each of the six 25-item scales constructed on the basis of the analysis described above possessed sufficient internal consistency to merit the title of "scale." The diagonal elements of Table 1 show the Kuder-Richardson formula 20 internal consistency reliabilities for each of the six scales. These reliabilities vary from a low of .90 to a high of .95, indicating the six adjectival

TABLE 1  
RELIABILITIES (DIAGONAL), CORRELATIONS (UPPER RIGHT) AND PARTIAL CORRELATIONS  
(LOWER LEFT) AMONG SEMANTIC SCALES

Scale	E+	E-	A+	A-	P+	P-
E+	.94	.03	.65	.23	.30	.50
E-	-.76	.95	.22	.63	.57	.39
A+	.33	-.50	.93	.08	.59	.32
A-	-.40	.33	-.77	.90	.35	.64
P+	-.36	.15	.15	-.28	.92	.18
P-	.06	-.16	-.38	.32	-.70	.91

semantic differential scales possess excellent internal consistency properties. Thus, the correlations among the scales may be fruitfully examined.

The correlations among the six scales in the adjective semantic differential study are presented in the upper right portion of Table 1. The correlations among polar opposites are in all three cases close to zero, but positive in sign. Furthermore, all 12 correlations between subscales from one semantic area and subscales from another semantic area are larger in absolute magnitude than the three correlations representing polar opposites of the same scale. Thus, for example, the correlation between two subscales, A+ and E+, representing different semantic content, is relatively high (.65), while the correlation between two scales designed to measure the same content, E+ and E—, is low and positive (+.03), and in a direction opposite to that expected from the hypothesis of consistent responses to semantic content alone. Thus, these results provide no evidence favoring the idea that semantic space is bipolar when assessed in the adjectival form described in this study. The results parallel those described previously by Green and Goldfried (3).

The picture changes radically when the acquiescent response style is taken into account. If acquiescence, as well as semantic meaning, accounts for the patterns of correlations in the upper right portion of Table 1, then partialling out the effects of acquiescence from the correlations should yield a pattern of relationships which correspond to expectations based on a three-factor bipolar theory of responding. The total number of adjectives checked across all six scales (which, incidentally, yielded a coefficient alpha internal consistency of .79) was used as a measure of acquiescence; and partial correlations were computed among the scales holding constant the effects of acquiescence. These results are presented in the lower left part of Table 1. In contrast to the results presented in the upper right part of Table 1, the pattern of correlations in the lower left corresponds perfectly to what one would expect on the basis of semantic content responding—polar opposites correlate highly negatively and the absolute values of all three polar opposite correlations are larger in magnitude than the 12 cross-scale correlations. A more complete reversal of findings could hardly have been expected.

#### D. DISCUSSION

The results of this study provide dramatic support for the assumption of bipolarity made by Osgood *et al.* in their construction of the well-known semantic differential scales. The very high negative correlations found between polar opposites, when the response style of acquiescence is controlled, show that semantic space can indeed be considered bipolar. Fifty item scales con-



constructed from the two polar opposite scales by keying in the direction of consistent semantic meaning show Kuder-Richardson 20 internal consistencies of .92, .89, and .88 for evaluation, activity, and potency. Thus, the pessimistic conclusions of Green and Goldfried (3) must be reversed. It should be noted, however, that the partial correlations among polar opposites are not so high as to preclude completely some meaning-determined nonoppositional response tendencies. In particular, the polar-opposite partial correlations are only in the .70s, while the reliabilities of the scales would allow for high correlations. Thus the author concludes that semantic scales are only "approximately" bipolar.

It seems quite likely that complex concepts may more readily be judged nonoppositionally than particular instances of simple concepts. Thus the raw correlation among polar opposites may be found to be more in the positive direction for complex concepts than for simple concepts. For simple concepts raw polar opposite correlations may even be negative in sign without controlling for the effects of acquiescence. An experimental comparison relating the degree of negative correlation among polar opposites to concept complexity could evaluate the possibility that some portion of what has been here labeled as acquiescence may also represent stimulus-determined nonoppositional responding.

An interesting final question can be raised. Does acquiescence response style always represent an irrelevant response tendency in the semantic domain? If the connotative meaning of a concept in the environment is characterized by the kind of adjectives which may be applied to it—as when someone is asked in an interview whether he considers war to be cruel, vigorous, or powerful—then a person's responses may well be better predicted by an adjectival, acquiescence-ridden instrument than the more traditional bipolar scales; it may not pay to isolate acquiescence response tendencies from attitude assessments in such cases. The initial parts of this study demonstrated that unipolar scales possessing excellent internal consistencies can be constructed to meet such assessment goals.

#### E. SUMMARY

An adjective version of the semantic differential was constructed. Each of the two polar ends of each of the three semantic factors could be assessed by subjects' responses to 25 adjectives. An analysis of the intercorrelations among the six scales, computed across 100 concepts rated by 200 subjects, revealed large positive correlations among scales from different semantic domains and essentially zero correlations among polar-opposite scales. This pattern was interpreted as consistent with an acquiescence hypothesis. Partial correlations

among the six scales, holding constant the total number of adjectives checked by a subject, revealed a correlational pattern highly consistent with the assumption that semantic space is approximately bipolar.

## REFERENCES

1. BENTLER, P. M. Multidimensional homogeneity scaling. Paper presented at Special Meeting, Psychometric Society, New York, N.Y., 1966.
2. CARROLL, J. B. Review of Osgood, Suci, and Tannenbaum, "The Measurement of Meaning." *Language*, 1959, **35**, 58-77.
3. GREEN, R. F., & GOLDFRIED, M. R. On the bipolarity of semantic space. *Psychol. Monog.*, 1965, **79**, No. 599.
4. HAMILTON, D. L. Personality attributes associated with extreme response style. *Psychol. Bull.*, 1968, **69**, 192-203.
5. HEISE, D. R. Semantic differential profiles for 1,000 most frequent English words. *Psychol. Monog.*, 1965, **79**, No. 601.
6. JENKINS, J. J., RUSSELL, W. A., & SUCI, G. J. An atlas of semantic differential profiles for 360 words. *Amer. J. Psychol.*, 1958, **71**, 688-699.
7. OSGOOD, C. E., SUCI, G. J., & TANNENBAUM, P. H. *The Measurement of Meaning*. Urbana, Ill.: Univ. Illinois Press, 1957.

*Department of Psychology*  
*University of California, Los Angeles*  
*Los Angeles, California 90024*

## HISTORICAL AND CONTEMPORARY ATTITUDES TOWARD NUMBERS\*<sup>1</sup>

*Department of Psychology, State University of New York, College at Brockport*

---

M. S. LINDAUER

---

The emotional meaning of numbers, as it relates to psychology, was examined in an historical context by Murphy (2). Because numbers have been found useful by both science and mysticism, the implication was drawn that equivocal if not contradictory attitudes characterize man's conception of numbers. To what degree has this conflict, inherent in the ambivalent qualities of numbers, been resolved in favor of either a positive or negative evaluation of numbers and concepts related to numbers?

In order to answer this question in its broadest terms, 138 quotations in 14 standard reference collections were selected by the author on the basis of their evaluative reference to numbers. This sample of commentary reflects the most interesting and perceptive thinking, stated at various periods in history, on the subject of numbers. "The genius, wit and spirit of a nation are discovered in its proverbs," said Francis Bacon in *Essays*. The notable sayings were drawn from 32 subject-listings in the tables of contents and indices of the references (a Thesaurus and synonym dictionary provided relevant terms). Thus the material dealt explicitly, as well as implicitly, with the concept of number. About half of the quotes (49 per cent) eventually used in the study came from the following three categories: "number" (including "figures," "statistics," "measurement," "counting," etc.), "certainty" (including "doubt" and "uncertainty"), and "fact" (including "proof," "disproof," "knowledge," "probability," etc.).

Three psychologist-judges evaluated the relevancy of the materials to the theme of numbers, and then judged each quote as either "positive" or "negative" (or "undecided") in its attitude towards numbers. Judges were instructed to try to ignore feelings of right or wrong, good or bad, and agreement or disagreement toward the quotes. Each judge evaluated one-half of the quotes in two sessions. Indication of the reliability of the evaluations is

---

\* Received in the Editorial Office, Provincetown, Massachusetts, on September 19, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> Professor Gardner Murphy's critical reading of an earlier version of this paper is gratefully acknowledged.

the fact that there was no difference between the judgments in the two sessions ( $\chi^2 = .82$ ,  $df = 4$ ,  $p > .05$ ). Disagreements between the judges were unanimously resolved in discussion in all but 19 cases (14 per cent of the original pool). Of the 119 quotes judged relevant and unanimously evaluated by all judges, 77 (65 per cent) were rated as reflecting a negative evaluation of numbers; this difference between positive and negative judgments was significant ( $\chi^2 = 10.30$ ,  $df = 1$ ,  $p < .01$ ).<sup>2</sup>

There were dates associated with 110 (92 per cent) of the quotes (none of which was later than 1937), thereby permitting an examination of time trends. (When more than one date was attributed to the quote, only the earliest one was used.) The modern technological era has evidently led to a greater preoccupation with numbers, since a little over a third of all quotes (39 per cent) were made within the most recent 100 years, while about another third (34 per cent) were stated between 800 B.C. and the end of the 17th century. With respect to the relative differences between positive and negative quotes, a difference in favor of a negative evaluation is evident in Figure 1. Except for the earliest period, there has been a greater percentage of negative rather than positive quotes at each time period, with the most recent period showing about a 2:1 ratio in favor of a negative evaluation. A sign test indicates that these differences over time in the negative direction were significant ( $p = .035$ ). A rank-order correlation of  $-.56$  between time intervals and frequency of negative quotes suggests ( $p < .10 > .05$ ) that the incidence of negative attitudes towards numbers has been increasing with time.

If numbers can be said to invoke an emotional attitude, the recorded statements of the notable men of our civilization apparently emphasize the negative components more than the positive ones. However, unknown are the selective factors which were operative in recording the quotes of various periods and accumulating them in collections. The use of a wide variety of sources, which selected quotes for various purposes and by different criteria, hopefully minimized some of the inherent biases of the sample of materials used. In addition, it should be pointed out that most of the quotes were stated by literary rather than scientific figures; and that most of them refer to the objective rather than subjective uses (*cf.*, 1) to which numbers may be put. Thus, when the "tender-minded" refer to "tough-minded" material, the apparent predom-

---

<sup>2</sup> A list of the quotes used and their evaluations has been deposited with the National Auxiliary Publications Service. Order Document No. 00103 from ASIS National Auxiliary Publications Service, c/o CCM Information Sciences, Inc., 22 West 34th Street, New York, New York 10001. Remit in advance \$3.00 for photocopies or \$1.00 for microfiche.



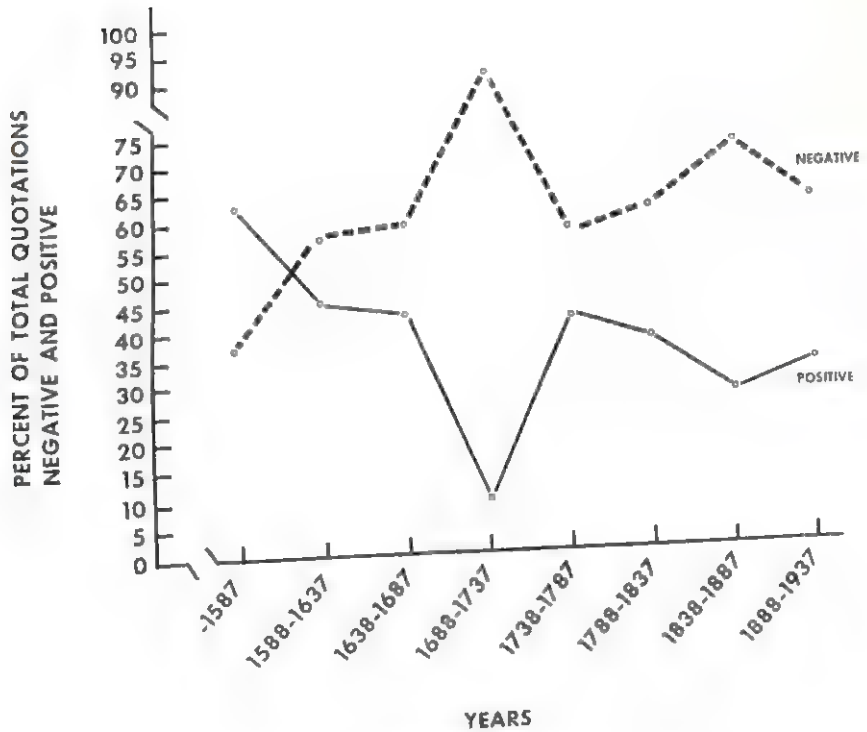


FIGURE 1  
PERCENTAGE OF POSITIVE AND NEGATIVE QUOTES RELATING TO NUMBERS (800 B.C.-1937)

inance of a negative evaluation of numbers is not surprising and should accordingly be cautiously interpreted.

#### REFERENCES

1. ANANT, S. S. Number theory. *Amer. Psychologist*, 1968, **23**, 204.
2. MURPHY, G. Pythagorean number theory and its implications for psychology. *Amer. Psychologist*, 1967, **22**, 423-431.

Department of Psychology  
State University College of New York  
at Brockport  
Brockport, New York 14420



## A COMPARISON OF PERSONALITY FACTOR STRUCTURE IN BRITISH AND AMERICAN UNIVERSITY STUDENTS\*<sup>1</sup>

*Department of Psychology, University of California, Los Angeles*

---

KAY JAMISON AND ANDREW L. COMREY

---

### A. INTRODUCTION

To determine the fundamental factors of human personality and how they can best be measured, the second author and his collaborators have pursued a personality research program (e.g., 4, 5, 8, 9, 15) based on the premise that for the purpose of locating major personality factors it is better to factor analyze total scores over collections of relatively homogeneous items, rather than to analyze single items. The factors which have emerged most clearly from this research have been named Shyness, Dependence, Empathy, Neuroticism, Compulsion, Hostility, and Socialization. The present investigation had two primary purposes: (a) using personality test materials designed to measure the above named factors, to compare the factor structures between a British and an American sample; and (b) to determine the differences in mean scores for the two samples on these same personality factors.

Several investigators have undertaken cross-cultural research between Britain and the United States. R. B. Cattell has established through several studies (e.g., 2, 3, 16) a cross-cultural constancy on the 16 PF, while at the same time demonstrating significant differences between the populations studied: e.g., British and American, Japanese and American, Australian and American. The results obtained in Cattell's British-American study will be discussed later in this paper.

Investigations of personality factors and school achievement have been carried out by Butcher, Ainsworth, and Nesbitt (1); differences between the "gifted" in the two countries have been investigated by Wall (19); Turner (17) has studied competitiveness and social acceptance among American and English students; Tyler (18) has studied the differences in the interests of English and American school children; and Honkavaara (14) has compared the relation of color-and-form reactors between British and American uni-

---

\* Received in the Editorial Office, Provincetown, Massachusetts, on September 26, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This research was supported by U.S. Public Health Grant MH 13055-01 and by a grant from the University of California.

versity students. Some of their results are compared with the results of this study later in the article.

Gorer (12, 13), using the cultural anthropological methods of Margaret Mead, has written about the cultural differences between the British and American societies, concluding that "the belief in the identity or similarity of the English and Americans is the greatest stumbling block that exists to mutual understanding and collaboration between the two peoples" (12, p. 12). The personal adviser on American affairs to British Ambassadors, Mr. P. S. Rankine, has suggested the establishment in both countries of an Anglo-American Institute "to conduct research into the issues which cause misunderstanding between this country (Britain) and America, and stimulate the production of books and television programmes which would help the two nations to understand each other better" (London *Observer*, February 25, 1968). In order to get rid of basic misunderstandings it will perhaps be helpful to examine the differences that exist between the British and the Americans with respect to certain personality factors.

## B. METHOD

### 1. *The Samples*

A total of 482 subjects participated in the research project. The 250 volunteers in the American sample, 125 males and 125 females, comprised almost entirely university students. A few university employees were included in the sample. The 232 subjects in the British sample, 115 males and 117 females, were all university students. The British sample comprised 129 Scottish students, 85 English students, and 18 students from other countries. Each subject in both samples was promised an analysis of his results in order to encourage participation in the project and honesty in filling out the inventory.

### 2. *The Inventory*

The same inventory, a version of the Comrey Personality Scales, was administered to both samples.<sup>2</sup> The inventory consisted of 216 items designed to measure seven major factors: Shyness, Dependence, Empathy, Neuroticism, Compulsion, Hostility, and Socialization. Each of these factors was

---

<sup>2</sup> The following items have been deposited with the National Auxiliary Publications Service: Comrey Personality Scales Inventory and answer sheet, matrices of correlations among FHIDs, matrices of minimum residual factors, matrices of criterion I factors, and matrices of criterion II factors. Order document number 00130 from ASIS National Auxiliary Publications Service, c/o CCM Information Sciences, Inc., 22 West 34th Street, New York, New York 10001, remitting in advance \$1.00 for microfiche or \$3.00 for photocopies.



made up of six to nine factored homogeneous item dimensions or FHIDs (+). A FHID consists of several items which were developed to measure the same variable and which were subsequently shown to be homogeneous through actual data analysis. Most of these items for the FHIDs had been used in prior investigations, but there were some new ones being tested in this study. Fifty-two FHIDs made up the Comrey Personality Scales used for this research project.

In order to avoid some of the problems of forced-choice and true-false-item formats, two response scales were used which had nine alternatives each. Scale X had the following choices: 1. Never, 2. Almost Never, 3. Rarely, 4. Occasionally, 5. Fairly Often, 6. Frequently, 7. Usually, 8. Almost Always, and 9. Always. Scale Y had these possible answers: 1. Absolutely Not, 2. Very Definitely Not, 3. Definitely Not, 4. Probably Not, 5. Possibly, 6. Probably, 7. Definitely, 8. Very Definitely, and 9. Absolutely. An X or a Y was put next to each item number on the inventory to suggest which scale was most likely to be suitable, but the subjects were informed that they could use whichever scale they felt to be more applicable.

In order to control for a possible acquiescence response set, approximately half of the 216 items were worded negatively with respect to the FHID name and half were worded positively. For example, two items on the FHID Seclusiveness were, "I try to avoid new contacts with people," and "At a party, I like to meet as many new people as I can." Like all items measuring the same FHID, they were well-separated in the inventory booklet. An earlier study (5) failed to establish any meaningful relationship between a measure of acquiescence and these factors, but it was thought best to control this possible source of bias as much as possible.

The 52 FHIDs for this analysis are listed below with a sample item for each one:

1. *Stage Fright*, 11Y. It would be hard for me to do anything in front of an audience.
2. *Follower Role*, 21X. I try to avoid positions of leadership.
3. *Seclusiveness*, 29Y. I try to avoid contacts with new people.
4. *Reserve*, 41X. In a group of people I keep quiet.
5. *Lack of Social Poise*, 51X. I find it difficult to talk with a person I have just met.
6. *Loss for Words*, 137X. In a group of people I find myself at a loss for words.
7. *Submission*, 43X. When someone puts pressure on me, I give in easily.

8. *Avoiding Personal Contact*, 112X. I have crossed the street to avoid having to speak to somebody I know.

9. *Passivity*, 114X. I try to avoid influencing other people.

10. *Need for Friends*, 104Y. It is very important for me to have close friends.

11. *Need for Love*, 84Y. I have a strong need for love and affection.

12. *Lack of Self Sufficiency*, 157Y. I am lost when left by myself.

13. *Succorance*, 7X. I depend on people to help me with my problems.

14. *Need for Approval*, 5Y. It is important for me to be accepted in my community.

15. *Need for Praise*, 8Y. If nobody praises my work, I find it hard to keep trying to do a good job.

16. *Need for Protection*, 87X. I try to avoid being in groups where the people say unkind things.

17. *Service*, 14Y. I have a strong desire to do something for the good of humanity.

18. *Helpfulness*, 54Y. I enjoy helping people even if I don't know them very well.

19. *Interest in People*, 158Y. I would like a job where dealing with people is the most important thing.

20. *Sympathy*, 33X. I am a very sympathetic person.

21. *Generosity*, 23X. I am generous with the poor.

22. *Kindness*, 132X. I try to avoid embarrassing anybody.

23. *Agitation*, 4X. My nerves seem to be on edge.

24. *Pessimism*, 16X. When I want something to happen, I have the feeling that it won't.

25. *Depression*, 78X. I feel that life is drudgery and boredom.

26. *Inferiority*, 150X. I have the feeling that the people I know are better than I am.

27. *Inadequacy*, 168X. I lack confidence in myself.

28. *Lack of Ego Strength*, 69X. The problems I face seem to be too big for me.

29. *Conformity*, 2X. I feel better doing what everyone else is doing.

30. *Intolerance of Nonconformity*, 129Y. Young people should be more willing than they are to do what their elders tell them to do.

31. *Acceptance of the Social Order*, 199Y. I would like the society in which I live to stay pretty much like it is now.

32. *Compliance*, 195X. I live up to the standards of conduct which are approved of by most people.

33. *Affiliation*, 102X. I avoid joining organizations if at all possible (item is reversed with respect to the FHID name).

34. *Respect for Law*, 85X. If a law is bad, you should obey it and try to get it changed rather than disobey it.

35. *Welfare of Loved Ones*, 175X. I feel obliged to do everything possible to see my family healthy and happy.

36. *Deference*, 47X. I dislike working under anybody (item reversed with respect to the FHID name).

37. *Social Activities*, 79X. I find social activities very enjoyable.

38. *Drive to Finish*, 18X. If I start a job, I finish it.

39. *Meticulousness*, 6X. I will go to great lengths to correct mistakes in my work which other people wouldn't even notice.

40. *Grooming*, 46X. I am very careful about my personal appearance.

41. *Routine*, 49X. I like to maintain a regular schedule of activities.

42. *Order*, 39X. I keep everything in its proper place so I know just where to find it.

43. *Cautiousness*, 140X. For me, the slow sure way is better than the fast, risky way.

44. *Belief in Human Worth*, 1Y. Most people are valuable human beings.

45. *Cynicism*, 9Y. Most public officials would accept bribes if they were large enough.

46. *Rhathymia*, 156Y. The best thing is to have fun now and not worry about the future.

47. *Defensiveness*, 91Y. You don't get far unless you are ready to fight off the competition.

48. *Psychopathy*, 169Y. If I could get rich legally by cheating people, I would do it.

49. *Vengeance*, 162Y. If I could cause a lot of trouble for one of my enemies without getting hurt myself, I would do it.

50. *Paranoia*, 143Y. There are times when everyone seems to be against you.

51. *Lack of Faith in Human Nature*, 115Y. Most people are out to get more than they give.

52. *Resentment*, 88Y. There are certain people I would particularly like to see put in their place.

### 3. The Analysis

Total scores for the FHIDs were obtained by summing the individual item scores over the groups of homogeneous items. The direction of negatively worded items (with respect to the FHID name) was reversed by subtracting

such item scores from 10. These total scores for the 52 FHIDs were then intercorrelated and the resulting  $52 \times 52$  matrix factor analyzed by the minimum residual method (Comrey & Ahumada, 1965).<sup>3</sup> Twenty-three minimum residual factors were extracted from the British data and 20 factors from the American data. Five of the British factors had no loadings of over .30; seven had only one variable appearing with a loading greater than .30. Because two of these residual factors had their one loading greater than .55, however, they were retained and rotated with the 11 major factors. Thus, 13 British factors were rotated by normal Criterion I of the Tandem Criteria for orthogonal analytic rotation (6). This criterion attempts to distribute as much variance as possible on a minimum number of factors, with the stipulation that variables appearing on a given factor must be correlated with one another. Of these 13 Criterion I factors, three had no FHIDs with loadings greater than or equal to .30; two others, shown here, were considered to be of minor importance (only FHIDs with loadings of .30 or more are listed): (a) IX. *Residual*. 47. Defensiveness, .30; 50. Paranoia, .30; and (b) XII. *Residual*. 38. Drive to Finish, .43. The remaining eight Criterion I factors, which were also the eight factors with the largest sums of squares of factor loadings, were rotated orthogonally by normal Criterion II of the Tandem Criteria, which distributes the variance more evenly among the factors to approximate a simple structure solution.

Of the 20 American minimum residual factors extracted, four had no loadings greater than .30 and eight had only one FHID with a loading greater than .30. The 14 minimum residual factors with the highest sums of squares of factor loadings were then rotated by Criterion I of the Tandem Criteria. Three of these Criterion I factors had no FHID with a loading greater than .30; three others, shown as follows, were considered to be of minor importance (only FHIDs with loadings of .30 or more are listed): (a) IX. *Residual*. 3. Seclusiveness, —.32; 37. Social Activities, .41; (b) X. *Residual*. 17. Service, .30; 52. Resentment, .32; and (c) XII. *Residual*. 41. Routine, .33. The remaining eight Criterion I factors, those with the largest sums of squares of factor loadings, were then rotated by Criterion II of the Tandem Criteria. Only eight Criterion I factors, therefore, appeared to be sufficiently important to rotate by Criterion II in both the British and American data.<sup>4</sup>

### C. RESULTS

The Criterion II eight factor solutions for the two sets of data are listed below. The British factor loadings are presented first and the American load-

<sup>3</sup> The computations were carried out on the IBM 7094 operated by the U.C.L.A. Campus Computing Network.

<sup>4</sup> See Footnote 2.



ings are in parentheses. If the loading was .30 or more in either analysis, the loadings for both groups are shown:

I. *Empathy*. 3. Seclusiveness, —.36 (.23); 5. Lack of Social Poise, —.35 (—14); 17. Service, .49 (.64); 18. Helpfulness, .70 (.77); 19. Interest in People, .66 (.61); 20. Sympathy, .59 (.64); 21. Generosity, .71 (.70); 22. Kindness, .54 (.61); 44. Belief in Human Worth, .32 (.44); 46. Rathymia, —.23 (—37); 47. Defensiveness, —.31 (—15); Psychopathy, —.48 (—44); 49. Vengeance, —.49 (—54).

II. *Submission*. 1. Stage Fright, .54 (.49); 2. Follower Role, .68 (.78); 3. Seclusiveness, .30 (.17); 4. Reserve, .57 (.44); 5. Lack of Social Poise, .68 (.33); 6. Loss for Words, .74 (.43); 7. Submission, .47 (.52); 8. Avoiding Personal Contact, .43 (.27); 9. Passivity, .61 (.67); 33. Affiliation, —.10 (—30); 36. Deference, .44 (.25); 43. Cautiousness, .24 (.33).

III. *Compulsion*. 14. Need for Approval, .33 (.19); 30. Intolerance of Nonconformity, .38 (.40); 34. Respect for Law, .31 (.32); 38. Drive to Finish, .41 (.60); 39. Meticulousness, .55 (.65); 40. Grooming, .52 (.54); 41. Routine, .73 (.50); 42. Order, .69 (.66); 43. Cautiousness, .52 (.32); 46. Rathymia, —.37 (—24).

IV. *Dependence*. 4. Reserve, —.31 (—24); 10. Need for Friends, .69 (.68); 11. Need for Love, .62 (.73); 12. Lack of Self Sufficiency, .46 (.42); 13. Succorance, .66 (.59); 14. Need of Approval, .39 (.27); 15. Need for Praise, .41 (.38); 16. Need for Protection, .55 (.48); 29. Conformity, .50 (.32); 37. Social Activities, .28 (.30).

V. *Shyness*. 3. Seclusiveness, .66 (.67); 4. Reserve, .40 (.62); 5. Lack of Social Poise, .27 (.75); 6. Loss for Words, .24 (.70); 8. Avoiding Personal Contact, .53 (.68); 12. Lack of Self Sufficiency, —.30 (—13); 29. Conformity, —.36 (.05); 30. Intolerance of Nonconformity, —.41 (—21); 33. Affiliation, —.63 (—17); 37. Social Activities, —.55 (—54); 40. Grooming, —.18 (—31); 44. Belief in Human Worth, —.40 (—14); 45. Cynicism, .45 (.05); 51. Lack of Faith in Human Nature, .58 (.13).

VI. *Neuroticism*. 7. Submission, .41 (.31); 15. Need for Praise, .40 (.30); 23. Agitation, .60 (.56); 24. Pessimism, .58 (.71); 25. Depression, .54 (.60); 26. Inferiority, .73 (.66); 27. Inadequacy, .76 (.79); 28. Lack of Ego Strength, .62 (.62); 38. Drive to Finish, —.28 (—30); 50. Paranoia, .38 (.36).

VII. *Hostility*. 8. Avoiding Personal Contact, .37 (.13); 17. Service, —.45 (—23); 18. Helpfulness, —.43 (.07); 22. Kindness, —.38 (—22); 24. Pessimism, .38 (.19); 44. Belief in Human Worth, —.57 (—51); 45. Cynicism, .68 (.72); 46. Rathymia, .31 (.31); 47. Defensiveness, .38 (.68); 48. Psychopathy, .44 (.50); 49. Vengeance, .60 (.44); 50. Paranoia, .68 (.68); 48. Psychopathy, .44 (.50); 49. Vengeance, .60 (.44); 50. Paranoia, .68 (.68);

.36 (.58); Lack of Faith in Human Nature, .57 (.59); 52. Resentment, .40 (.46).

VIII. *Socialization*. 12. Lack of Self Sufficiency, —.04 (.31); 14. Need for Approval, .47 (.69); 25. Depression, —.31 (—15); 29. Conformity, .29 (.74); 30. Intolerance of Nonconformity, .24 (.54); 31. Acceptance of Social Order, .67 (.57); 32. Compliance, .58 (.49); 33. Affiliation, .30 (.54); 34. Respect for Law, .45 (.33); 35. Welfare of Loved Ones, .46 (.27); 36. Deference, .43 (.50); 38. Drive to Finish, .30 (.04).

#### D. DISCUSSION

Dependence, Empathy, Neuroticism, Socialization, Compulsion, and Hostility emerged in this study much as predicted; Shyness, however, split into two factors (in both the British and American samples) which were designated as Shyness and Submission. Submission was defined by such FHIDs as Stage Fright, Follower Role, Submission, and Passivity; while such FHIDs as Reserve, Seclusiveness, and Avoiding Personal Contact grouped clearly together on the Shyness factor. Although there is a correlation between the two factors (approximately .60 for both the samples), it is clear that a person who is shy and reserved need not be submissive and passive and, likewise, a person who is submissive and passive need not be shy and reserved. The patterns of loadings for Submission and Shyness are basically the same for the two samples with the exception of two FHIDs—Lack of Social Poise and Loss for Words—which reversed their places on the two samples. That is, the British Lack of Social Poise and Loss for Words both have high loadings (.68, .74) on the Submission factor and low loadings (.27, .24) on Shyness; whereas for the American data, Lack of Social Poise and Loss for Words have relatively low loadings (.33, .43) on Submission, and high loadings (.75, .70) on Shyness. Another difference between the two samples on the Shyness factor is that three Hostility FHIDs showed up on the British sample (Belief in Human Worth, —.40; Cynicism, .45; Lack of Faith in Human Nature, .58) and not on the American.

As in the most recent study of this series (9), several of the Hostility FHIDs had loadings greater in absolute value than —.30 on the Empathy factor. This is not unreasonable as Empathy and Hostility have exhibited a correlation of —.44 in two previous studies (8, 9); the correlation between them in this investigation was approximately —.55 for the two sets of data.

Empathy, Dependence, and Neuroticism all exhibited a close match in factor structure between the two samples. The Compulsion factor likewise displayed a good fit, although two FHIDs—Drive to Finish and Meticulous-

ness—showed up on the American sample with higher loadings than on the British sample and two other FHIDs—Routine and Cautiousness—had higher loadings in the British data than in the American data. The most significant difference was for the Cautiousness FHID, .52 in the British solution and .32 in the American solution. Cautiousness did not show up with a loading greater than .35 anywhere in the American results.

Socialization showed the least correspondence of factor structures. On the whole, the American results were much better (e.g., Conformity: .29 on British, .74 on American; Intolerance of Nonconformity: .24 on British, .54 on American). Socialization is a newly added factor in the Comrey system of personality factors, hence not as much research has been done on it as for the other factors; it is perhaps not surprising, therefore, that Socialization fails to match in the two samples as well as the other factors.

To assess the correlation between the British and American factor structures, the loadings on each of the eight British factors were intercorrelated with the loadings on each of the American factors. The following correlations between the corresponding British and American factors were obtained: Empathy, .94; Submission, .90; Compulsion, .91; Dependence, .95; Shyness, .68; Neuroticism, .97; Hostility, .86; and Socialization, .60. It should be pointed out that the degree of agreement attained was through applying the same analytic criteria to the two sets of data independently. No effort was made to maximize the degree of agreement between the two solutions.

The *t* ratios for the differences in mean factor and FHID scores between the British and American samples are shown in Tables 1 and 2. It can be seen that the British students were, on the average, more shy, submissive, compulsive, and hostile, and less dependent than their American counterparts. There was little difference between the two groups on the Empathy, Socialization, and Neuroticism factors. Examining the differences in means on FHID scores, however, makes the pattern of differences a bit clearer. For instance, although there was not much difference between the samples in Empathy, Neuroticism, and Socialization, there were several differences in the individual FHIDs that made up these factors. The British had significantly lower scores on the Conformity FHID and significantly higher scores on the Acceptance of Social Order FHID; both of these FHIDs are part of the Socialization factor. Likewise, the British had significantly lower scores on Agitation, Depression, and Inferiority but significantly higher scores on Pessimism and Inadequacy; these FHIDs are part of the Neuroticism factor. On the Empathy factor the Americans had significantly higher scores on Service, while the British had significantly higher scores on Generosity.

TABLE 1  
FHID SCORE MEANS AND STANDARD DEVIATIONS IN BRITISH AND AMERICAN SAMPLES

FHID name	British		American		<i>t</i> ratio
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
1. Stage Fright	23.2	6.8	19.8	7.1	5.5
2. Follower Role	22.0	5.7	21.5	5.8	0.8
3. Seclusiveness	19.6	6.3	16.6	6.0	5.4
4. Reserve	22.7	6.0	17.0	5.9	10.5
5. Lack Social Poise	16.7	5.6	18.6	6.1	3.4
6. Loss for Words	16.6	6.0	15.8	6.3	1.4
7. Submission	13.1	5.2	16.0	5.0	6.4
8. Avoiding Pers. Con.	21.3	5.5	15.3	5.0	12.3
9. Passivity	21.2	5.0	18.9	5.4	5.0
10. Need for Friends	11.9	2.9	13.7	2.9	6.9
11. Need for Love	13.6	2.8	13.9	2.9	1.2
12. Lack Self Suffic.	14.8	6.8	22.2	7.7	11.1
13. Succorance	20.2	7.0	23.0	6.7	4.4
14. Need for Approval	43.1	10.4	44.6	11.0	1.6
15. Need for Praise	15.3	5.3	18.5	5.4	6.7
16. Need for Protection	23.3	4.5	25.4	4.5	5.1
17. Service	15.6	4.6	17.9	4.9	5.2
18. Helpfulness	36.6	6.9	36.6	7.0	0.1
19. Interest in People	39.6	6.8	39.8	6.8	0.4
20. Sympathy	19.6	3.7	19.8	3.6	0.7
21. Generosity	31.2	6.7	27.4	7.0	6.0
22. Kindness	34.2	4.8	34.4	4.6	0.4
23. Agitation	12.6	5.2	16.2	5.6	7.3
24. Pessimism	19.0	5.4	15.3	5.4	7.5
25. Depression	9.4	4.2	12.1	4.4	7.0
26. Inferiority	12.7	6.1	15.0	6.2	4.0
27. Inadequacy	17.7	5.1	14.8	5.0	6.4
28. Lack Ego Strength	6.9	2.5	6.8	2.6	0.2
29. Conformity	17.6	4.7	20.0	5.2	5.4
30. Intol. Nonconf.	9.9	3.0	10.7	3.0	2.9
31. Accep. Social Order	11.7	3.1	8.6	3.3	10.6
32. Compliance	28.9	4.1	25.6	4.3	8.7
33. Affiliation	10.2	3.3	10.8	3.5	2.0
34. Respect for Law	14.8	3.3	12.5	3.3	7.1
35. Welfare Loved Ones	15.7	2.7	13.6	3.4	7.5
36. Deference	16.0	4.4	13.2	4.4	7.0
37. Social Activities	25.0	5.3	25.5	5.4	1.0
38. Drive to Finish	29.1	4.4	25.9	4.7	7.8
39. Meticulousness	28.4	6.6	27.2	6.5	2.0
40. Grooming	28.3	5.0	25.4	5.3	6.2
41. Routine	25.6	5.9	20.0	5.6	10.8
42. Order	26.0	6.2	21.9	6.4	7.2
43. Cautiousness	37.8	8.2	32.6	8.2	6.8
44. Belief Human Worth	30.1	6.8	35.9	5.8	10.1
45. Cynicism	44.9	9.6	37.1	7.6	9.8
46. Rhythymia	28.6	7.0	36.7	7.3	12.4
47. Defensiveness	24.5	6.2	29.9	5.8	10.0
48. Psychopathy	14.9	6.7	16.2	7.0	2.1
49. Vengeance	14.3	5.3	13.4	5.0	2.0
50. Paranoia	17.2	3.3	16.8	3.8	1.1



TABLE 1 (*continued*)

FHID name	British		American		<i>t</i> ratio
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
51. Lack Faith Human Nature	26.3	5.2	21.0	4.1	12.4
52. Resentment	22.4	5.5	22.0	5.5	0.8

TABLE 2  
FACTOR SCORE MEANS AND STANDARD DEVIATIONS IN BRITISH AND AMERICAN SAMPLES

Factor name	British		American		<i>t</i> ratio
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
1. Socialization	92.5	18.9	94.8	20.8	1.3
2. Shyness	86.2	22.9	77.8	23.9	3.9
3. Submission	102.2	19.7	93.1	21.5	4.8
4. Neuroticism	78.2	21.6	80.2	22.8	1.0
5. Compulsion	137.4	20.5	120.3	20.4	9.2
6. Dependence	84.4	15.8	94.6	15.1	7.3
7. Empathy	176.8	26.8	175.8	27.1	0.4
8. Hostility	150.3	26.6	143.1	24.5	3.1

There were significant differences between the mean factor scores on Shyness and Submission; but, again, an examination of the differences between FHID scores illustrates the nature of the disparity; the British had higher scores on Seclusiveness, Reserve, and Avoiding Personal Contact.

The Americans had higher scores on all of the FHIDs comprising the Dependence factor, and the British had higher scores on all of the FHIDs comprising the Compulsion factor. On the Hostility factor the British had higher scores on those FHIDs measuring primarily cynicism—e.g., Cynicism, Lack of Faith in Human Nature—and a lower score on Belief in Human Worth.

It is interesting to compare these results with those obtained in another cross-cultural study of a similar nature (3). They compared the scores of American and British students on the Sixteen Personality Factor Questionnaire and discovered that "... on the whole, the British students were less anxious, more introverted, more sensitive and more radical; the American students were more anxious, more extraverted, less sensitive and more conservative" (3, p. 15). They also found that the British group showed significantly higher ego strength than the American group. The results obtained here tend to support their findings that British students are less anxious and more introverted and, insofar as Conformity is a measure of conservatism, that British students are more radical than American students. Somewhat at odds with this finding, however, is the fact that the British students in the present

investigation were significantly higher on the FHID Acceptance of the Social Order than were the American students.

Gorer's (12) findings that Americans tend to reject authority, that they are far more ready than the British to give and receive friendship, and that Americans have a dread of loneliness and isolation seem to be supported by the findings in this study that the British score far higher on the Respect for Law and Deference FHIDs, lower on the Need for Friends FHID, and far higher on the Reserve and Seclusiveness FHIDs. Honkavaara (14), in his study of the color-and-form test, concluded that the English student was "more intellectual, sensitive, shy, and individualistic; the American is practical and socially conforming." Our results tend to support his findings that the British are more shy and nonconforming, although again it must be pointed out that the British scored higher on the Acceptance of Social Order FHID, but lower on the Conformity FHID than did the Americans. The complex question arises here as to how one can differentiate between conformity with respect to one's personal life and conformity with respect to the acceptance of the society within which one lives.

More sophistication is needed to aid in the interpretation of results from cross-cultural research studies. For example, is one to account for differences in mean scores on the basis of genuine differences on the factor being measured or in terms of idiomatic usage, cultural differences in attitudes toward the taking of personality tests or other determinants? Fisher (11) suggests that the differences obtained from administering the Jesness Inventory to English and American institutionalized boys "reflects not the personality differences indicated but differences in test-taking techniques related to a relatively pronounced acquiescent response set among the English boys." Two other researchers, Mott in England and Jesness in Sacramento, disagree with Fisher's skepticism about the international "exchange ability" of psychological tests, but no definitive work has been done to support either side.<sup>5</sup>

### E. CONCLUSIONS

The major conclusions suggested by this study are as follows: First, the personality factors measured by the Comrey Personality Scales emerged, at least in the American sample, substantially as predicted from the results of prior studies in this series. Second, there was a substantial correspondence between the British and American personality factor structures based on variables derived from the Comrey Personality Scales. This reaffirms a similar finding in an earlier comparison of American and Italian subjects (10). Third, there were many significant differences in mean FHID and factor

<sup>5</sup> Quoted from personal correspondence from Fisher to the first author.

scores between the two samples, a conclusion in agreement with the findings of other investigators in the cross-cultural field. The British were found to be more shy, submissive, compulsive, and hostile, and less dependent than their American counterparts.

## REFERENCES

1. BUTCHER, H. J., AINSWORTH, M., & NESBITT, J. E. Personality factors and school achievement: A comparison of British and American children. *Brit. J. Educ. Psychol.*, 1963, **33**, 276-285.
2. CATTELL, R. B. A cross-cultural check on second stratum personality factor structure: Notably of anxiety and exvia. *Aust. J. Psychol.*, 1965, **17**, 12-23.
3. CATTELL, R. B., & WARBURTON, F. W. A cross-cultural comparison of patterns of extraversion and anxiety. *Brit. J. Psychol.*, 1961, **52**, 3-16.
4. COMREY, A. L. Factored homogeneous item dimensions in personality research. *Educ. Psychol. Meas.*, 1961, **21**, 417-431.
5. ———. Personality factors Compulsion, Dependence, Hostility, and Neuroticism. *Educ. Psychol. Meas.*, 1964, **24**, 75-84.
6. ———. Tandem criteria for analytic rotation in factor analysis. *Psychometrika*, 1967, **32**, 143-154.
7. COMREY, A. L., & AHUMADA, A. Note and Fortran IV program for minimum residual factor analysis. *Psychol. Rep.*, 1965, **17**, 446.
8. COMREY, A. L., & JAMISON, K. Verification of six personality factors. *Educ. Psychol. Meas.*, 1966, **26**, 945-953.
9. COMREY, A. L., JAMISON, K., & KING, N. Integration of two personality factor systems. *Multivariate Behav. Res.*, 1968, **3**, 147-160.
10. COMREY, A. L., MESCHIERI, L., MISITI, R., & NENCINI, R. A comparison of personality factor structure in American and Italian subjects. *J. Personal. & Soc. Psychol.*, 1965, **1**, 257-261.
11. FISHER, R. M. Acquiescent response set, the Jesness Inventory, and implications for the use of "foreign" psychological tests. *Brit. J. Soc. & Clin. Psychol.*, 1967, **6**, 1-10.
12. GORER, G. *The American People*. New York: Vail-Ballou Press, 1948.
13. ———. *Exploring English Character*. New York: Criterion Books, 1962.
14. HONKAVAARA, S. Comparison of the relation of color- and form-reactors at Harvard and London University. *J. of Psychol.*, 1958, **46**, 23-24.
15. JAMISON, K., & COMREY, A. L. Further study of dependence as a personality factor. *Psychol. Rep.*, 1968, **22**, 239-242.
16. TSUJIOKA, B., & CATTELL, R. B. Constancy and difference in personality structure and mean profile, in the questionnaire medium, from applying the 16 P.F. test in America and Japan. *Brit. J. Soc. & Clin. Psychol.*, 1965, **4**(4), 287-297.
17. TURNER, R. H. Preoccupation with competitiveness and social acceptance among American and English students. *Sociometry*, 1960, **23**, 307-325.
18. TYLER, L. E. A comparison of the interests of English and American school children. *J. Genet. Psychol.*, 1956, **88**, 175-181.
19. WALL, W. D. Highly intelligent children: I. The psychology of the gifted. *Educ. Res.*, 1960, **2**, 101-110.

Department of Psychology  
University of California, Los Angeles  
Los Angeles, California 90024





## RELATION BETWEEN INTERNAL-EXTERNAL CONTROL AND AGGRESSION\*

*Department of Psychology, University of Manitoba*

CHRISTENE BLANTON WILLIAMS AND FLORENCE E. VANTRESS

### A. INTRODUCTION

The internal-external control (I-E) dimension developed by Rotter (7) refers to the degree to which an individual believes that his reinforcements are contingent upon his own behavior. The internally oriented individual believes that the reinforcements follow as a consequence of his own behavior; while, conversely, the externally oriented individual believes that reinforcements are controlled by forces independent of his behavior (i.e., fate, luck, chance, or other individuals). A scale (I-E Scale) has been devised to measure internal-external control and has been used in a number of recent studies (4, 8).

Among the factors not yet explored through the I-E Scale are those of hostility and aggression. There are reasons to believe that externally oriented individuals would be more aggressive and more hostile than internally oriented individuals. First, Rotter, Seeman and Liverant (9) found a significant relationship between external control and authoritarianism. The fact that authoritarianism and hostility are positively related has been confirmed by several studies (5, 10). It follows, then, that external control might be related to hostility also. Second, the frustration aggression hypotheses (1, 3) suggest that frustration often leads to aggression. In the case of the external scorers, frustration and subsequent aggression can be expected for several reasons. The external scorer does not perceive himself as able to manipulate the environment in such a manner as to obtain desired goals. According to Minton (6), the externally oriented individual sees himself as powerless. Hence goal blocking and frustration are possible through many forms of bad luck and chance occurrences. In addition, Tolor and Janowiec (11) found that high external scorers perceived that their mothers manifested authoritarian control and hostile rejecting tendencies; thus, through their perception of being controlled and rejected, it is likely that frustration has occurred and could lead to future aggressive behavior.

\* Received in the Editorial Office, Provincetown, Massachusetts, on September 30, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

The purpose of this paper is to attempt to determine the relationship between scores on the I-E Scale and scores on a measure of aggression, the Buss Durkee Hostility Inventory (2). This particular inventory was selected, since it measures several different types of aggression, and thus gives a relatively detailed picture of the relationship between internal-external control and aggression.

## B. METHOD

### 1. *Subjects*

The subjects were 235 introductory psychology students (114 males and 121 females) at the University of Manitoba who volunteered for this investigation as partial fulfillment of course requirements.

### 2. *Scales*

The material consisted of Rotter's (8) Internal-External Scale and the Buss-Durkee Hostility Inventory (2). The I-E Scale consists of 29 forced-choice items and is scored in terms of the total number of external choices selected. A high score represents an external expectancy, and a low score an internal expectancy of reinforcement. Nationally obtained norms (8) were used as the criteria for internal or external scorers instead of making a classification in terms of the present sample distribution of scores: internally oriented individuals were defined as those scoring 6 or below on the I-E Scale, and the externally oriented individuals were defined as those scoring 10 or above on that scale.

The Buss-Durkee Hostility Inventory (BD) is a 75-item, self-report questionnaire. It consists of eight subscales designed to tap several aspects of the construct "hostility-aggression." Factor analysis of the subscales (1) results in two clusters of scales: Hostility (the Suspicion and Resentment subscales) and Aggression (the Assaultive, Indirect, Negativism, and Verbal subscales).

## C. RESULTS

The correlation between scores on the I-E Scale and scores on the BD scale was small but significant ( $r = .27$ ,  $p < .05$ ). Therefore, the more one sees his reinforcements as contingent upon factors other than his own behavior, the more aggression and hostility he reports.

In order that the relationship between internal-external control and aggression might be further examined, internal scorers and external scorers were compared on the eight BD subscales. The externally oriented individuals scored significantly higher than the internally oriented individuals on the

Resentment ( $F = 17.65, p < .001$ ), Verbal ( $F = 3.93, p < .05$ ), Suspicion ( $F = 16.38, p < .001$ ), Indirect ( $F = 8.91, p < .01$ ), and Irritability ( $F = 17.65, p < .001$ ) subscales of the BD. There were no differences between the two groups on the remaining three subscales: Negativism ( $F = 0.67, p > .05$ ), Guilt ( $F = 0.79, p > .05$ ), and Assault ( $F = 2.11, p > .05$ ).

#### D. SUMMARY

Rotter's I-E Scale and the Buss-Durkee Hostility Inventory were administered to 235 college students. The results indicate that there was a small positive correlation between internal-external control and aggression. In addition, external scorers scored significantly higher than internal scorers on five of the eight subscales of the Buss-Durkee Hostility Inventory.

#### REFERENCES

1. BUSS, A. *The Psychology of Aggression*. New York: Wiley, 1961.
2. BUSS, A., & DURKEE, A. An inventory for assessing different kinds of hostility. *J. Consult. Psychol.*, 1957, 21, 343-348.
3. DOLLARD, J., DOOB, L., MILLER, N., MOWRER, O., & SEARS, R. *Frustration and Aggression*. New Haven, Conn.: Yale Univ. Press, 1939.
4. LEFCOURT, H. M. Internal versus external control of reinforcement. *Psychol. Bull.*, 1966, 65, 206-210.
5. MEER, S. J. Authoritarian attitudes and dreams. *J. Abn. & Soc. Psychol.*, 1955, 51, 74-78.
6. MINTON, H. L. Power as a personality construct. In B. A. Maher (Ed.), *Progress in Experimental Personality Research*. New York: Academic Press, 1967. Pp. 229-267.
7. ROTTER, J. B. *Social Learning and Clinical Psychology*. Englewood Cliffs, N.J.: Prentice-Hall, 1954.
8. ———. Generalized expectancies for internal vs. external control of reinforcement. *Psychol. Monog.*, 1966, 80(1).
9. ROTTER, J. B., SEEMAN, M., & LIVERANT, S. Internal versus external control of reinforcement: A major variable in behavior and theory. In N. F. Washburne (Ed.), *Decisions, Values, and Groups* (Vol. 2). London, England: Pergamon Press, 1962. Pp. 473-516.
10. SIEGEL, S. M. The relationship of hostility to authoritarianism. *J. Abn. & Soc. Psychol.*, 1956, 52, 368-372.
11. TOLOR, A., & JANOWIEC, J. E. Body boundary, parental attitudes and internal-external expectancy. *J. Consult. & Clin. Psychol.*, 1968, 32(2), 206-209.

Department of Psychology  
The University of Manitoba  
Winnipeg, Canada





## EXPERIMENTER EXPECTANCY EFFECTS IN THREE DISSIMILAR TASKS\*

*Saint Louis University*

---

RICHARD L. WESSLER

---

### A. PROBLEM

Rosenthal's (5) finding that the expectancies of experimenters (*E*s) about the performance of their subjects (*S*s) can affect *S*s' performance has serious implications for the design and conduct of research. But the experimental evidence for *E* expectancy effects is based primarily on a single task, the ratings of photographs of faces on a scale of perceived success-failure. While this task has been a useful one for studying various aspects of the *E* expectancy phenomenon, the results from a single task cannot indicate the variety of tasks and situations to which the findings may be generalized, nor whether one should anticipate *E* expectancy effects in data produced by human *S*s on other tasks. Supportive evidence from other investigations has been mixed and is based on tasks of the same type as photo-ratings: *S*s made judgments or reported perceptions for which there were no objectively correct or incorrect responses (1, 2, 3, 6).

The present study asked whether *E*s who affected *S*s' responses on the photo-rating task would also affect *S*s' performances on dissimilar tasks. A taxonomy described elsewhere (7) was followed, and the tasks were chosen for their varying degrees of obviousness of correct response. The photo-rating task calls for judgments and *S*s do not know which of the judgments they might make is best—unlike a timed psychomotor task in which the fastest response is obviously best. Somewhere between these two poles lie tasks in which *S*s' judgments are sought but for which there are objectively correct responses (e.g., estimating lengths of lines). The susceptibility of judgments to the influence of others is well known. When judgments are called for, *S* is more dependent upon *E* for cues about what response will be most positively evaluated by *E*—i.e., what is the most appropriate response to make—than when *S*s merely must demonstrate some ability (4). The overall hypothesis is that the more obvious the "correct" response is to *S*s from the

---

\* Received in the Editorial Office, Provincetown, Massachusetts, on October 7, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

nature of the task, the less susceptible is *Ss*' behavior to the influence of *Es*' expectancies.

## B. METHODS

### 1. *Tasks*

Three tasks were employed: the rating of photos of faces, the judging of line lengths, and the tapping of dots into circles. The photo-rating task consisted of 10 photographs of faces chosen from 30 photos clipped from magazines and rated by judges on a 20-point scale of success-failure. The 10 photos whose mean scores and sum of mean scores deviated least from zero were retained for use in the experiment. These scaling procedures conformed as closely as possible to those described by Rosenthal (5). *S*'s task was to rate each photo on a 20-point scale how much success or failure the person is experiencing.

The line-judging task consisted of 12 pairs of cards, on one card of which a single vertical line was displayed; the other card contained three lines of varying lengths, one of which was the same length as the criterion line on the other card. *S*'s task was to state which line was the same length as the criterion line.

The tapping task, patterned after a subtest of the MacQuarrie Mechanical Abilities Test, consisted of 10 rows of 10 circles each on a single sheet of paper. The *S*'s task was to tap three dots into as many circles as possible during a 45-second time period.

### 2. *Subjects*

*Es* were 11 seniors and six graduate students enrolled in the writer's research methods course. None admitted in a postexperimental debriefing session to having awareness of Rosenthal's work. *Ss* were 18 males and 22 females, primarily freshmen and sophomores, drawn from introductory sociology courses. Each *E* ran from one to four *Ss*.

### 3. *Design*

The experimental variables in this study were created by leading *Es* to expect certain performances from *Ss* by means of written instructions. For the photo-rating task, *Es* read either that "previous research has shown that subjects should obtain scores of about +5" or "about -5." For line judging, *Es* read either that "subjects usually overestimate the length of lines" or that they underestimate them. For the tapping task, *Es* read either that *Ss* should "complete about 100 circles" (fast performance) or "about 50 circles"

(slow performance) in the prescribed time limit. All that *Es* were to say to *Ss* was contained in their instruction sheets.

#### 4. Procedures

*E* was given the self-explanatory instructions, data recording sheets, photo and line cards, pencils, and a stopwatch, by an assistant of the writer. Both *Es* and *Ss* were told that the purpose of the research was to standardize the results of the tests employed. *E* was allowed time to acquaint himself with the procedures, but not to ask questions of the assistant. He then saw his *Ss* and administered the tests in the order described above. Expectations were not changed from *S* to *S*; rather *E* maintained his orientation for each *S*.

#### C. RESULTS AND DISCUSSION

For the photo-rating task the mean rating for *Ss* run under the +5 condition was  $-.40$ , while the mean rating for *Ss* run under the  $-5$  condition was  $+.15$ . The results are opposite those that had been predicted; therefore, a one-tailed test of these data is not appropriate. An analysis of variance was performed, but revealed that no main effect or interaction approached significance.

Line-judging scores were obtained by subtracting for each *S* the number of underestimations of length (short errors) from the number of overestimations (long errors). The means for the overestimation group and the underestimation group were 6.2 and 4.5 respectively. Although these means are in the predicted direction, an analysis of variance reveals that no main effect or interaction was significant.

Scores on the tapping task were obtained by counting the number of circles completed within the time limit. For the fast expectancy group,  $\bar{X} = 71.4$ ; and for the slow expectancy group,  $\bar{X} = 69.8$ . The difference, though in the predicted direction, is not significant when subjected to an analysis of variance.

Since, according to these analyses, *E* expectancy effects were not obtained for any of these tasks, additional analyses were performed with the use of a more liberal criterion for expectancy effects. *Ss* whose mean photo-rating scores had the same algebraic sign as that of the expectancy given their *Es* were considered to have been "biased." Twenty of the 40 *Ss* were so designated.

Of the 20 "biased" *Ss*, 10 were run by *Es* in the long-error expectancy group and 10 by short-error expectancy group *Es*. An examination of the quality of error made by each group reveals that for the line-judging, the

ratio of long errors to all errors was 30:32 for the long-error group, and 17:32 for the short-error group. Disproportionately more of the errors made by the long-error group were long errors ( $\chi^2 = 3.03$ ,  $df = 1$ ,  $p < .05$ , one-tailed). However, when the "biased" *Ss*' scores on the tapping task were compared, no significant differences were found. The mean for the 10 *Ss* in the fast performance group was 71.2, and 69.6 for the slow performance expectancy group. These trends are congruent with the hypothesis that the more obvious the correct response is to *S*, the less susceptible *S*'s performance is to *E* expectancy effects, probably because *S* need not seek information about how his performance will be evaluated by *E*.

These findings also suggest that the failure of some studies to obtain *E* effects when using other tasks may not be due simply to the resistance of the task, but may be due to the failure of the investigator to create conditions under which *E* expectancy effects can be obtained on the most susceptible task (photo-rating). Thus, the prescreening of *E* and *S* combinations for magnitude of effects is suggested when dealing with questions of their generalization.

#### D. SUMMARY

*Ss* whose ratings of photographs of faces on a scale of perceived success-failure approximated their *Es*' expectancies also produced judgments of line lengths similar to their *Es*' expectancies. However, no differences in speed of hand tapping could be attributed to their experimenters' expectancies. With regard to the generalization of experimenter expectancy effects across tasks, these results suggest that when a desirable performance level (e.g., speed) is obvious to *S* from the nature of the task, his performance is less susceptible to such effects.

#### REFERENCES

1. MCFALL, R. M. Unintentional communication: The effect of congruence and incongruence between subject and experimenter constructions. Doctoral dissertation, Ohio State University, Ann Arbor, Mich.: Univ. Microfilms, 1965, No. 66-1809.
2. MARWIT, S. J., & MARCIA, J. R. Tester-bias and response to projective instruments. *J. Consult. Psychol.*, 1967, **31**, 253-258.
3. MASLING, J. Differential indoctrination of examiners and Rorschach responses. *J. Consult. Psychol.*, 1965, **29**, 198-201.
4. RIECKEN, H. W. A program for research on experiments in social psychology. In N. F. Washburne (Ed.), *Decisions, Values and Groups (Vol. II)*. New York: Pergamon Press, 1962. Pp. 25-41.
5. ROSENTHAL, R. Experimenter effects in behavioral research. New York: Appleton-Century-Crofts, 1966.

6. STRAUSS, M. E. Examiner expectancy: Effects on Rorschach experience balance. *J. Consult. & Clin. Psychol.*, 1968, **32**, 125-129.
7. WESSLER, R. L. Experimenter expectancy effects in psychomotor performance. *Percept. & Motor Skills*, 1968, **26**, 911-917.

*Department of Sociology and Anthropology*  
*Saint Louis University*  
*Saint Louis, Missouri 63103*





## RETROACTIVE INHIBITION OF CONNECTED DISCOURSE AS A FUNCTION OF MEANINGFULNESS AND OF DEGREE OF ORIGINAL LEARNING\*<sup>1</sup>

*Department of Psychology, University of Saskatchewan, Saskatoon, Canada*

---

JOHN A. MILLS AND GORDON WINOCUR<sup>2</sup>

---

### A. INTRODUCTION

It has been suggested (5) that retroactive inhibition (RI) would not be found when highly meaningful material is used. It is also possible to argue that retroactive facilitation (RF) might be produced if meaningfulness is high enough. Underwood and Schulz (11) have demonstrated that the various measures of meaningfulness, including the frequency of occurrence of verbal units (e.g., nonsense syllables, words), intercorrelate strongly. Now, consider a case in which connected discourse is being used in an RI paradigm with two levels of meaningfulness and with the level of meaningfulness equivalent in original learning (OL) and interpolated learning (IL). The supposition would be that, in sentences consisting of high frequency words, it is very likely that many of the words would be associates of the others. If OL and IL are similar, then associations tending to maintain the coherence of the lists come from two rather than one source. At the same time, because of the ease of learning, the differentiation (9) between the lists is high. Therefore recall should be better in the experimental condition than in the control condition. With an equivalent degree of similarity in the low meaningful condition, an interpolated list should serve to disrupt the relatively fragile "structure" (3) of the original list, so that RI should occur.

Better support for the argument appears to come from a study by King (2). King used verbal material of differing degrees of contextual constraint (so that meaningfulness, in effect, was varied) in an RI design and found that, whatever the level of constraint of OL, the interpolation of fifth order material or of a text produced RF in recall, whereas the interpolation of first or third order material produced RI. Because King used the method

---

\* Received in the Editorial Office, Provincetown, Massachusetts, on October 8, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This research was sponsored by a grant from the President's Fund of the University of Saskatchewan.

<sup>2</sup> The authors are indebted to Mrs. Lorraine Smith, who assisted in the collection of the data.

of whole presentation and free recall it was impossible for him to control degree of learning. In the absence of such control, it is possible to argue that the relatively high degree of learning which must have occurred in interpolated material with high levels of constraint would increase the amount of differentiation of the OL from the IL. The IL would then merely preserve the Ss' set so that King's finding could be the result of distributed practice.

Similar arguments could be used to explain other findings that apparently support Mills and Kessel's argument (5, 6). In these studies, connected discourse of a type that would probably be rated high in meaningfulness was used. The small amounts of interference obtained may not have been a function of level of meaningfulness as such, but of the high degrees of learning characteristic of such material (10). It is known that, with nonsense lists, the amount of RI starts to decline once the degree of learning in OL becomes considerable (4), a phenomenon which should be even more marked when ordinary prose is used.

These issues were settled by varying both meaningfulness and degree of learning in OL. Since only these two factors were of interest, similarity between OL and IL was held constant at zero. The serial anticipation method, which allows considerable control to be exerted over degree of learning, was used.

## B. PROCEDURE

### 1. *Subjects*

The Ss were 90 female university students, all under 25 years old, enrolled in first year psychology classes at the University of Saskatchewan.

### 2. *Materials*

Meaningfulness was varied by the making up of 24-word English sentences containing words of differing degrees of frequency of occurrence as determined by the Thorndike and Lorge (8) general word count. One pair of sentences was designated high frequency (HF) and consisted of 23 AA and one A words. The next pair was designated medium frequency (MF) and contained either eight or nine AA words and two A words; the remaining words varied in frequency from 20 to 46 times per million. Finally, a low frequency (LF) pair had seven or eight AA words and either one or zero A words; the other words varied in frequency from one to 39 times per million (either 11 or 12 of these had frequencies of 19 times per million or less). Because Glanzer (1) has found that "content" words (nouns, verbs, adjectives, and adverbs) have different characteristics during learning from "func-

tional" words (prepositions, pronouns, and conjunctions), the number of the latter was held constant at four. In addition, the mean number of articles (a, an, the) was held relatively constant across frequency levels. The means per sentence were: LF: 2.0; MF: 2.5; HF: 2.0.

The sentences were as follows:

LF 1: "Billboards fulfill the dual role of obstructing the surrounding landscape from perplexed, bewildered tourists and maintaining an adequate camouflage for concealing turnpike patrol vehicles. . . ."

LF 2: "Beehive honeycombs demonstrate unusual intricate geometrical frameworks which enable the structure to withstand maximal applied transverse stress when related to its bulk and texture. . . ."

MF 1: "Largely, the dispute was concerning whether hunters have truly developed suitable garments to resist freezing in the terribly severe climate of the remote provinces. . . ."

MF 2: "In the election debates the previous candidate for mayor advocated that confusion concerning traffic projects illustrated the lack of keen professional ability on council. . . ."

HF 1: "The young children, coming from poor homes, would like playing in wide beautiful parks where they could laugh, shout, and ride on nice swings. . . ."

HF 2: "Everybody said they were getting very hungry so big tables were put under shady trees and food was brought out as quickly as possible."

The sentences were presented on Gerbrands memory drums with all punctuation marks inserted. The symbol "000" marked the beginning of each sentence.

### 3. *Design*

A random groups design was used with three levels of meaningfulness, two conditions of learning (presence or absence of the interpolated list), and two levels of OL. At each level of meaningfulness in OL, the equivalent level of meaningfulness was assigned in IL. Each sentence was used with almost equal frequency in OL and IL.

### 4. *Method*

The level of OL was varied as follows. First, 60 Ss (10 Ss per condition) were run to a criterion of one errorless trial in OL, IL, and relearning (RL). The protocols were then examined and it was decided that if Ss were run in OL until they had made between three and five errors (79.1 per cent to 87.5 per cent learning), then it would be possible to equate approximately the level

of OL across the level of meaningfulness. A further 30 Ss were then run to this criterion in OL and to a criterion of one errorless trial in both IL and RL. In the high OL condition, each sentence was used five times in OL and IL; in the low OL condition each sentence was used either twice or three times in OL and IL.

The length of the interpolated interval was 20 minutes. For the control group, the whole of this was filled by a pursuit rotor task; for the experimental group, any part of the interval not occupied in IL was used on the same task. OL was preceded by two trials of number guessing, and RL by one. The number series 123, 456, 789, 123, etc. was used. The usual instructions for rote learning by the serial anticipation method were given. The drum was set at a three-second rate; the intertrial interval was six seconds.

## C. RESULTS

### 1. *Acquisition*

It was important to determine that the sentences used differed in meaningfulness between frequency levels but were equivalent within frequency levels. The criterion used was rate of learning in the high OL condition. The mean trials to criterion were as follows: LF 1: 11.4; LF 2: 12.3; MF 1: 6.9; MF 2: 7.5; HF 1: 3.8; HF 2: 3.4. A one-way analysis of variance of trials to criterion gave an  $F$  of 7.53 ( $df$  5,54;  $p < .001$ ). A Duncan's Range test showed that both the HF sentences were significantly different from both the LF sentences. MF 1 was significantly different from LF 2; there were no other significant differences involving MF sentences. At the same time, the differences within frequency levels were not significant and, since the means fell in the expected rank order, it was concluded that the sentences had fulfilled the criteria.

In the low OL condition, it was inevitable that some Ss would have error scores that would not fall within the limits which had been set. A danger here was that a relatively large number of Ss would be rejected from the HF condition because they would make two errors or less on Trial 2, which would mean that this condition would contain a relatively high proportion of slow learners. However, the number of rejected Ss was roughly the same across levels of meaningfulness (five from LF, three from MF, and seven from HF). A  $\chi^2$  test showed that these differences were not significant ( $\chi^2 = 1.079$ ,  $df$  2).

### 2. *Retention*

The means and standard deviations of the correct responses on Trial 1 of RL, together with the amounts of relative RI are shown in Table 1. A fac-



torial analysis of variance was performed on these scores, the results of which are shown in Table 2. The only significant main effects were those involving the conditions of learning ( $F = 14.11$ ,  $df$  1,78,  $p < .001$ ) and degree of OL ( $F = 87.97$ ,  $df$  1,78,  $p < .001$ ). There was a significant interaction between these two variables ( $F = 12.94$ ,  $df$  1,78,  $p < .001$ ). In order to explore these effects more fully, separate analyses of variance were performed within the high OL and low OL conditions. In the former case, no significant effects were found. In the latter, the only significant effect was due to the conditions of learning ( $F = 11.72$ ,  $df$  1,24,  $p < .01$ ). It was therefore concluded that significant RI had occurred only when degree of OL was low, that there was no RI when degree of OL was high, and that the meaningfulness of the material had no effect on the amount of RI obtained. Such RF as occurred was not statistically significant.

In the high OL condition, RI was dissipated very rapidly. For the experimental condition, mean trials to relearn for the LF, MF, and HF sentences

TABLE 1  
MEANS AND STANDARD DEVIATIONS OF CORRECT RESPONSES ON  
TRIAL 1 OF RL, WITH AMOUNTS OF RELATIVE RI

Level of learning	Condition of learning	Meaningless		
		LF	MF	HF
Low	Control	$18.6 \pm 3.05$	$19.6 \pm 2.51$	$21.2 \pm 1.48$
	Experimental	$15.8 \pm 3.56$	$14.6 \pm 5.13$	$16.6 \pm 2.97$
	Relative RI (per cent)	15.0	25.5	21.6
High	Control	$22.9 \pm 1.29$	$22.6 \pm 1.26$	$22.5 \pm 1.51$
	Experimental	$21.9 \pm 1.91$	$21.4 \pm 1.65$	$23.0 \pm 1.25$
	Relative RI (per cent)	4.4	5.3	+ 2.2*

\* The + sign indicates RF.

TABLE 2  
ANALYSIS OF VARIANCE OF TOTAL ERRORS ON TRIAL 1 OF RL

Source	<i>df</i>	<i>MS</i>	<i>F</i>
Between	11	55.181	
Between Meaningfulness (A)	2	9.700	1.973
Between Conditions (B)	1	69.344	14.105***
Between OL Level (C)	1	432.449	87.967***
A $\times$ B	2	3.145	.639
A $\times$ C	2	3.255	.662
B $\times$ C	1	63.606	12.938***
A $\times$ B $\times$ C	2	4.606	.936
Within	78	4.916	
Total	89	11.129	

\*\*\*  $p < .001$ .

were 1.9, 1.4, and .5 respectively. The corresponding values for the control condition were 1.0, .9, and 1.1. Because of imperfect learning in OL, it took longer for Ss to reach criterion in the low OL condition. For the experimental condition, mean trials to relearn for the LF, MF, and HF conditions were 2.8, 3.4, and 2.4 respectively. The corresponding values for the control condition were 2.8, 2.8, and 1.2.

The percentage of errors which were intrusions in each of the experimental conditions is shown in Table 3. The great majority of the intrusions were intralist errors. Because, with one exception, there were no statistically significant differences between control and experimental conditions, the error data for these conditions were pooled.  $\chi^2$  tests showed that, at both levels of OL, there was a statistically significant tendency for the proportion of intrusions to increase as meaningfulness rose. There were no statistically significant differences in the proportion of intrusions between levels of OL within frequency levels.

TABLE 3  
PERCENTAGE OF INTRUSIONS DURING RL

Level of OL	Condition of learning	Meaningfulness		
		LF	MF	HF
Low	Control	47	56	100
	Experimental	46	68	64
High	Control	40	45	68
	Experimental	33	52	80

#### D. DISCUSSION

The results of the study are very clear-cut. Significant RI was found only when level of OL was low, no matter how high the meaningfulness of the material. This establishes that the low amounts of interference reported previously (5, 6) were a function of high levels of learning in OL. The bearing of the findings on the study by King (2) should also be considered. A major difference between the present experiment and King's study is that he used what he calls "low control" procedures (free learning, unlimited recall time, etc.), while we used what he calls "tight control" procedures (the standard serial anticipation method). However, he discounts this difference as a possible source of explanation for success or failure to obtain RI with connected discourse. If this position is granted, then, in one of the conditions of the present experiment (that in which high meaningful material was used with a high degree of learning in OL), King should predict RF. But statistically significant RF was not found, nor were any of the interaction terms involving the

factor of meaningfulness significant. The relative facilitation in RL in the high OL HF condition reinforces the hypothesis which has already been made that RF is due to performance factors operating during learning.

The fall in the omission: intrusion ratio as a positive function of meaningfulness gives some indication of the mechanism producing RI. Slamecka (7) has already noted that, in connected discourse, the better the recall, the more likely it is that an error will be an intrusion. Since the associative strength of high meaningful material at the end of learning is greater than that of low meaningful material, the rising proportion of intrusions is not unexpected. There is an apparent contradiction, however, in that, at both levels of learning, amounts of RI were equivalent across levels of meaningfulness. An apparent inference from this finding is that the proportion of omissions should remain constant as meaningfulness varies. The contradiction can be resolved by making the hypothesis that although equivalent amounts of interference are produced at each level of meaningfulness, the manifestations of the interference differ. With low meaningful material, the relatively low associative strength induces a high proportion of omissions. With high meaningful material, the relatively high associative strength makes it likely that a word from within the sentence (or an associate of such a word) will be given. But because the structure of the sentence has been disrupted by IL, the words will be emitted in the wrong order. To this extent, but to this extent alone, there is conformity with the hypothesis that guided earlier studies (5, 6).

#### E. SUMMARY

An experiment employing the RI paradigm, in which the material was connected discourse, was designed. The two major variables were degree of OL and level of meaningfulness. Meaningfulness was varied by generating sentences containing words of differing frequencies of occurrence. The main findings were that RI was found only with a low degree of OL and that amount of RI was equivalent across levels of meaningfulness. It was concluded that the frequent failures to find RI with connected discourse are due to the relatively high degrees of learning which are a characteristic of this material.

#### REFERENCES

1. GLANZER, M. Grammatical category: A rote learning and word association analysis. *J. Verb. Learn. & Verb. Behav.*, 1962, **1**, 31-41.
2. KING, D. J. Contextual constraint and retroaction in verbal material. *J. Gen. Psychol.*, 1966, **75**, 39-63.
3. MANDLER, G. From association to structure. *Psychol. Rev.*, 1962, **69**, 415-427.
4. MELTON, A. W., & IRWIN, J. MCQ. The influence of interpolated learning on retro-

- active inhibition and the overt transfer of specific responses. *Amer. J. Psychol.*, 1940, **53**, 173-203.
5. MILLS, J. A., & KESSEL, F. Retroactive inhibition of descriptive prose. *Psychol. Rep.*, 1965, **17**, 917-918.
  6. MILLS, J. A., & SACKS, S. Proactive inhibition of descriptive prose as a function of the length of retention interval. *J. Gen. Psychol.*, 1967, **76**, 241-249.
  7. SLAMECKA, N. J. Retroactive inhibition of connected discourse as a function of practice level. *J. Exper. Psychol.*, 1960, **59**, 104-108.
  8. THORNDIKE, E. L., & LORGE, I. *The Teacher's Word Book of 30,000 Words*. New York: Teachers Coll., Columbia Univ., 1944.
  9. UNDERWOOD, B. J. The effect of successive interpolations on retroactive and proactive inhibition. *Psychol. Monog.*, 1945, **59**, Whole No. 273.
  10. ———. Degree of learning and the measurement of forgetting. *J. Verb. Learn. & Verb. Behav.*, 1964, **3**, 112-129.
  11. UNDERWOOD, B. J., & SCHULZ, R. W. *Meaningfulness and Verbal Learning*. New York: Lippincott, 1960.

*Department of Psychology*  
*University of Saskatchewan*  
*Saskatoon, Canada*

## PERFORMANCE CHARACTERISTICS OF FIELD DEPENDENT AND INDEPENDENT INDIVIDUALS ON AN AUDITORY SIGNAL DETECTION TASK\*

*Department of Psychology, St. John's University, Jamaica, New York*

VICTOR J. DEFazio AND WILLIAM F. MORONEY<sup>1</sup>

### A. INTRODUCTION

The cognitive control of field articulation, originally tested and conceived by Witkin (10) as visual independence and later as field independence (12), is concerned with the individual's ability to overcome competing cues so as to perceive or extract an item from the perceptual field in which it appears with a greater or lesser dependence on the field. Cognitive controls—i.e., field-dependence-independence, leveling, sharpening, etc.—are conceived of as slowly changing, developmentally stabilized structures which are relatively invariant over a given class of situations and intentions. These operate despite shifts in situational and behavioral contexts typical of cognitive activity from moment to moment (2). A series of studies (11) has tended to establish field independence or the analytical approach firmly within this definition.

Sherman (7) has criticized the analytic field approach, especially in terms of well established sex differences. Sherman claimed that the key measures of the analytic cognitive approach are substantially related to space perception. She concluded the term "analytical" itself implied an unwarranted generality, since the construct appeared unrelated to certain areas: e.g., verbal performance (11). Sherman also noted that "the link between sex, sex roles, and spacial skill could account for a considerable part of the relationship between personality variables and performance on the perceptual tasks. It would also account for the fact such correlations are higher with the spacial measures (12). The impressive stability of field independence measures could also occur because they are related to sex typing which itself is very stable" (7, p. 298).

If skills in spacial tasks were learned to a considerable extent and if opportunities for learning were sex typed, then, for example, feminine males would be less likely to engage in male sex-typed activities which would develop such

\* Received in the Editorial Office, Provincetown, Massachusetts, on October 14, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> Now at Naval Aerospace Medical Center, Pensacola, Florida 32512.



skills. Thus, generally speaking, male sex activities would have promoted field independence (FI), and female sex typing would have promoted field dependence (FD). Individuals low in the analytic field approach (FD) did in fact show characteristics more typical of women than men (11).

The rod and frame test (RFT) and the embedded figures test (EFT) have been utilized as the primary methods of measuring field articulation. The EFT requires an individual to recognize and trace a simple geometric figure in a larger, complex, and often colored design. This test has been significantly correlated with a number of personality variables, including achievement, motivation, dependency, aggressiveness, interest in people, etc. (5). Consistent sex differences have been found with this measure of field articulation: viz., males were FI and females were FD (10).

In order to delimit clearly the term "analytic approach," it seemed desirable to investigate the existence of such an approach in other sense modalities with the use of different tasks. With few exceptions (1, 4, 8), studies of field articulation have used visual tasks as the dependent variable. Therefore, investigation of the auditory modality might prove useful in testing the generalizability of this concept.

Jackson's investigation, which used an auditory detection task under traditional psychophysical procedures, was noteworthy. He required the subjects, classified on the basis of the EFT, to detect a word against a background of noise. A significant correlation between the EFT and the performance score was found for men ( $p < .05$ ) but not for women. However, Witkin, noting the distinction between embedding and distracting factors, classified the experimental procedures used by Jackson as a distracting task (i.e., one which required the subject to extract the stimulus—a word—from a relatively disorganized field). The development of the theory of signal detection (TSD) procedures may negate this objection.

The detection of an auditory signal in a field of white noise seemed a perfectly analogous task to the EFT. The following features were common to both EFT and TSD procedures: (a) an observer was required to pick out or extract a predetermined signal or item from a field in which the signal had been incorporated; (b) the *S* was informed of the probability of a signal's presence; (c) neither task relied heavily on memory. Price (6) provided the rational and empirical evidence for the use of TSD procedures in personality and perception research.

If a field analytic approach, as proposed by Witkin *et al.* (12), is a permanent, stabilized, and pervasive style of response, then it should enable an observer to respond with the same consistent style over a variety of tasks in dif-

ferent modalities. This is in agreement with Witkin's position that response style is concerned with a central interpretation of incoming information. With the use of an auditory signal detection task—TSD procedure (9)—it was expected that significant differences would be noted between the FI and FD groups, as well as between male and female in total response time ( $T$ ), criterion values ( $B$ ), and sensitivity ( $d'$ ).

## B. METHOD

### 1. Subjects

The paid volunteer  $S$ s were 10 male and 10 female graduate and undergraduate students ranging in age from 19 to 47 years.

### 2. Apparatus and Procedure

A binaurally presented 1000 cps tone, produced by a Spectral Dynamics sweep oscillator (Model SD 104-5), was used as the stimulus. The output of the oscillator was monitored by a Berkeley EPUT meter (Model 554) and attenuated by a General Radio microvolter (Model 546C). A Grason-Stadler electronic switch (Model 829E) was used to mix and control the onset and cessation of the auditory stimuli. The stimulus was masked by white noise (intensity =  $-4$  db with a reference of 1.5 volts) produced by a Grason-Stadler white noise generator (Model 901B). The output of the white noise generator was monitored by an EICO VTVM (Model 255). A Koss PRO-4 headset was used for the mechanical translation of the auditory stimuli.

A logic program which both determined the occurrence of a signal plus noise (SN) trial or a noise alone (N) trial and provided a means for recording the subject's hits (SN-y), misses (SN-n), false alarms (N-y), and correct rejections (N-n) was assembled. A Behavioral Research Systems (BRS) precision probability unit (PP-I) was used to determine the probability of the SN condition. A BRS binary counter (BC-I) was used to limit each session to 200 forced choice trials. Four Sedeco counters recorded the  $S$ 's decision as SN-y, SN-n, N-y, or N-n. The total running time for the experiment was recorded by a Hunter KlocKounter. A response box which allowed the subject to determine the onset of each trial and report the presence or absence of a signal was provided.

Taped instructions were used to explain both the task involved and the procedure to the subject.  $S$ s were familiarized with both the tone and white noise and trained until their numbers of hits and misses were approximately equal. Subjects were advised that the *a priori* probability of a signal occur-

rence was .50. During the experimental session subjects responded to three blocks (spaced one minute apart) of 200 trials each.

The short form of the EFT (3) was individually administered to the Ss. On the basis of their EFT scores Ss were divided into FD ( $N = 10$ ) and FI ( $N = 10$ ) groups. On the basis of  $t$  tests significant differences were noted between FD and FI groups ( $t = 1.76, p < .05$ ) and between males and females ( $t = 6.02, p < .01$ ).

Upon completion of the experimental session three analyses of variance (one-way classification) were performed on T, B, and  $d'$ . Male dependent, male independent, female dependent, and female independent constituted the levels of each analysis. An orthogonal comparisons procedure was applied to the results. Correlations between EFT scores and T, B, and  $d'$  were determined.

### C. RESULTS AND DISCUSSION

The present study utilized a task analogous to the EFT (i.e., a task which required the disembedding of a stimulus, except that the auditory modality was involved. No significant difference in T, B, or  $d'$  scores was noted between the FI or FD groups or between sexes ( $F_s < 1.00$ ). The results of the orthogonal comparisons procedure also failed to attain significant  $F$  values. No significant correlations were obtained between EFT scores and T ( $\rho = .175, p > .05$ ), B ( $\rho = .023, p > .05$ ) and  $d'$  ( $\rho = .170, p > .05$ ).

#### 1. *Sensitivity and Criterion*

Previous investigations into the relationship between EFT scores and sensitivity have revealed a relationship between EFT score and both pain tolerance (8) and audition (4). However, the investigations cited above used traditional psychophysical methods which, as Swets (9) has indicated, should no longer be considered adequate measures of sensitivity, since they are contaminated by criterion judgments. In the present study, the TSD approach permitted the differentiation of sensitivity ( $d'$ ) and criterion (B). With this method no significant differences in  $d'$  and B were noted between the FD and FI groups, nor was any significant correlation noted between EFT score and  $d'$  or B. This lack of either a significant difference or relationship indicated that while individuals classified as FD or FI may exhibit greater or lesser sensitivity, in the performance of a visual task, this characteristic does not influence their performance on an auditory task. In light of this finding, a TSD investigation employing a visual detection task is contemplated as a means for determining the applicability of the term sensitivity in investigations with a field-analytic orientation.

## 2. Time

A field dependent person has been described as reacting more to the field than to internal direction. Indeed, the field dependent subjects employed in this investigation took significantly longer to detect and report the required visual patterns contained in the EFT backgrounds than did the FI subjects. If the concept of field articulation is to be applicable across modalities, then the field dependent individuals should also have taken longer to detect and report the presence or absence of the auditory signal contained in the background than field independent subjects. However, in this investigation no differences were noted in terms of the time required for decision making on the auditory task. It may therefore be concluded that the term field articulation can only be used in relation to those tasks involving visual or visual motor tasks. The lack of significance supports Sherman's contention that the concept of "field analytic" approach has been erroneously generalized to denote a central process. In fact, the literature tends to reinforce this position. Witkin *et al.* (12) reported a significant relationship between visual motor tasks and field articulation, whereas no significant relationship was noted between field articulation and another nonvisual motor task: i.e., verbal learning (11).

## D. SUMMARY

Sherman (7) reported that the "field-analytic approach" was erroneously generalized to include all sense modalities. The Embedded Figures Test (EFT), which requires Ss to detect a figure in a background was used to differentiate the field-dependent group ( $N = 10$ ) from the field-independent group ( $N = 10$ ). In an analogous task, the theory of signal detection procedures (TSD) was used, and Ss were required to detect a tone in a white-noise background. No significant differences were noted between groups with respect to response time, criterion values, or sensitivity; nor was any relationship noted between EFT score and performance on the auditory detection task. These findings lend support to Sherman's contention.

## REFERENCES

1. AXELROD, S., & COHEN, L. Senescence and embedded figures performance in vision and touch. *Percept. & Motor Skills*, 1961, 12, 283-288.
2. GARDNER, R., HOLZMAN, P. S., KLEIN, G. S., LINTON, H., & SPENCE, D. P. Cognitive control. *Psychol. Iss.*, 1959, 1, Whole No. 4.
3. JACKSON, D. N. A short form of Witkin's embedded figures test. *J. Abn. & Soc. Psychol.*, 1956, 53, 254-255.
4. ———. Stability in resistance to field forces. Unpublished Doctor's dissertation, Purdue University, Lafayette, Indiana, 1965.
5. KATO, W. A fundamental study of the RFT. *Japanese Psychol. Res.*, 1965, 7, 61-68.

6. PRICE, R. H. Signal detection in personality and perception. *Psychol. Bull.*, 1966, 66, 55-62.
7. SHERMAN, J. A. Problem of sex differences in space perception as aspects of intellectual functioning. *Psychol. Rev.*, 1967, 74(4), 290-299.
8. SWEENEY, D. R., & FINE, B. J. Pain reactivity and field independence. *Percept. & Motor Skills*, 1965, 21, 757-758.
9. SWETS, J. A. Signal Detection and Recognition by Human Observers. New York: Wiley, 1964.
10. WITKIN, H. A. The nature and importance of individual differences in perception. *J. Personal.*, 1949, 18, 145-170.
11. WITKIN, H. A., DYK, R. B., FATERSON, H. F., GOODENOUGH, D. R., & KARP, S. A. Psychological Differentiation. New York: Wiley, 1962.
12. WITKIN, H. A., LEWIS, H. B., HERTZMAN, M., MACHOVER, K., MEISSER, P. B., & WAPNER, S. Personality through Perception. New York: Harper, 1954.

*Department of Psychology*  
*St. John's University*  
*Jamaica, New York 11432*



## THE EFFECT OF LEADING CONTOUR ON THE RELATIVE LENGTHS OF MOVING LIGHT AND DARK ARCS\*<sup>1</sup>

*Department of Psychology, The University of New England, Australia*

GORDON STANLEY AND ROBERT JACKSON

### A. INTRODUCTION

Ansbacher (1) observed that when a illuminated arc line on a dark surround is rotated around a central fixation point at a constant speed, it appears much shorter than its stationary length. He reported that the amount of shrinkage was positively related to the speed of rotation between zero and 75 rpm. Marshall and Stanley (3), using viewing conditions similar to Ansbacher except for the use of reflected instead of transmitted light and a greater peripheral viewing angle (20 degrees), found substantially less shrinkage in the light arc than Ansbacher and noted that a dark arc on light surround did not shrink when rotated. In a further study using conditions of increased similarity to Ansbacher, Stanley (6) confirmed that light arcs contract under rotation whereas dark arcs do not. In both these studies there was a relatively small tendency for the dark arc to elongate.

Stanley (8) reported that when an arc line with red and green sections is rotated around a central fixation point, Ss report seeing red followed by yellow when the leading half of the arc is red and the trailing half green, and see green followed by yellow when the order of color is reversed. He also found that the apparent length of the arc line is shorter when rotating than when stationary. An explanation of this effect may be given with the use of the notion of the successive excitation and inhibition of adjacent receptor units. On the assumption that a funnelling process (2, p. 628) occurs over a certain time interval, the regions being excited within this time interval may pool and summate with the region of maximal excitation, in this case the point of initial stimulation. The funnelling of the points of lesser excitation into the leading "on" region would create a sensory distribution of light much shorter than the physical distribution, which would cause the arc to be seen as shorter than its physical length. In support of this interpretation, Stanley (7) found, using a

\* Received in the Editorial Office, Provincetown, Massachusetts, on October 21, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This research was supported by a research grant from the University of New England to the first author.

50 cm diameter, dark disc with an illuminated arc line and a visual angle of 12 degrees between the central point of fixation and the arc line, that the brightness of the arc line was enhanced with the increasing speed of rotation between zero and 75 rpm.

The foregoing interpretation of previous research stresses the importance of the leading edge for the appearance of the shrinkage effect, and the experiment to be reported was designed to investigate the effect of the presence or absence of a leading contour on apparent shrinkage. Sherrington (4), using a half dark and half light disc with a dark arc jutting from the dark half and a light arc jutting from the light half, found that when the disc was rotated in such a way that the jutting "teeth" preceded the semicircles of the same color (i.e., no trailing contour), the ringband associated with the teeth had a higher critical flicker frequency (cff) than the rest of the disc. When rotated in the opposite direction (i.e., no leading contour), the cff for all points of the disc was approximately the same. The present study used discs similar to Sherrington and attempted to determine whether the difference in contour associated with the two directions of rotation affected an *S*'s comparison of the lengths of the light and dark teeth (arcs).

The apparatus for the present experiment was a modified color wheel; and as such wheels subtend a smaller visual angle than has previously been used in Ansbacher research, peripheral angle of viewing was also varied. To enable this variation, the *S* was required to view the disc either by fixating its centre (subtending a visual angle of seven degrees 10 minutes to the arc lines) or by viewing from either of two fixation points subtending visual angles of 15 degrees 42 minutes and 31 degrees 24 minutes respectively.

## B. METHOD

### 1. *Subjects*

The *Ss* were 10 male and 10 female undergraduates from the University of New England. All had normal vision or vision corrected to 20/20 and all were naive with respect to the Ansbacher effect.

### 2. *Apparatus*

The apparatus consisted of a modified color wheel of 16 cm radius rotated by a Singer sewing machine motor, its speed controlled by gears and the voltage input controlled by a Variac. The disc was rotated at a constant speed of 90 rpm, this being checked occasionally by a tachometer and found to be reasonably constant. There were three fixation lights located in a radial path of

100 cm from the *S*'s seating position so that a green light was located in the centre of the disc, a red light at 15 degrees 42 minutes, and a yellow light at 31 degrees 24 minutes, respectively, from the disc centre. The central fixation point was obtained by the use of a partial mirror and was only visible when *S* looked straight ahead. The *S* was seated in a swivel chair and looked at a dark wall (luminance = 25 ft lamberts) illuminated by a 40 watt incandescent lamp between trials and, on instruction, swivelled his chair around so that he faced the appropriate fixation point.

The stimuli consisted of two sets of three discs, each disc being half black and half white. On one set of the discs a black arc line 12 cm long and .8 cm wide was located at a radial distance of 12.5 cm from the disc centre, the arc being positioned so that it jutted out of the black sector. A white arc line was correspondingly positioned so that it jutted out of the white sector. The white arc line was varied in length for each disc, being either 8, 12, or 16 cm long. The other three discs were similarly constructed, the difference between the two sets being that, when rotated in a clockwise direction, the arc lines on the first set preceded the sectors of the same color; whereas the arc lines on the second set of discs followed the sectors of the same color. This meant that the effect produced by the second set of discs was the same as rotating the first set in a counterclockwise direction.

The discs were screwed onto the color wheel by means of a brass screw and were changed manually for each trial. During the experiment, the discs were illuminated (luminance = 50 ft lamberts) by a 100 watt incandescent lamp which was turned on for a two second interval by a Venner timer.

### 3. *Experimental Design*

The 20 *S*s were tested individually and randomly assigned to one of the two conditions. In the first condition, 10 *S*s were tested with the use of the three discs where, on clockwise rotation, the jutting arcs preceded the sectors of the same color. In the second condition, 10 *S*s were tested with the use of the other three discs where, on clockwise rotation, the jutting arc lines followed the sectors of the same color. For both conditions the disc was rotating at a constant speed of 90 rpm. Testing was completed over three sessions, in most cases on three consecutive days, each session being conducted at the same time of day. In any one session, each *S* made judgments of the relative lengths of the arcs for every stimulus at each fixation point. For a given session the nine conditions were each presented seven times in a randomly determined order except for a run restriction of two. The task was for the *S* to judge which of the two simultaneously-presented arcs was longer.

#### 4. Procedure

*S* was seated in a swivel chair and instructed to look at the lighted wall except when *E* told him to fixate one of the three colored lights. *S* was shown a miniature example of the stimulus arrangement and told that on any given trial the black arc could be longer than the white arc or *vice versa*. He was then told to judge which arc line was longer and was requested to place an eyepatch over his left eye.

#### C. RESULTS

Figure 1 shows the percentage of times the white arc lines were judged longer than the 12 cm, black arc line for each condition of rotation. By inspection it can be seen that, except for the straight-ahead fixation position with no leading contour, all conditions show a tendency for less white longer judgments, suggesting the presence of the Ansbacher shrinkage effect. Clearly presence of the leading contour results in many fewer white longer judgments under the straight-ahead fixation position. As peripheral angle is increased to 15 degrees 42 minutes, the presence or absence of a leading contour produces little difference in the relative judgments of arc length. At 31 degrees 24 minutes there is a slight tendency for the absence of the leading contour to result in more white longer judgments. From inspection of the left-hand side of Figure 1 it can be seen that in these two conditions of greater peripheral angle of viewing there is an increase in "false alarm" responses (i.e., an increase in the number of times the eight cm white arc is judged to be longer than the 12 cm black arc). This, together with the flattening of the curves, suggests that discriminability deteriorates as peripheral angle of viewing increases.

To check on the significance of the effect of the presence or absence of the leading contour in the straight-ahead position, the number of white longer judgments was obtained for each subject in the two groups. A *t* test for the significance of the difference between the means of the two groups yielded a ratio of 4.94 ( $df = 18$ ,  $p < .01$ ).

#### D. DISCUSSION

The present results suggest that under the straight-ahead fixation position the presence of a leading contour has a marked effect on the relative lengths of light and dark arcs, its absence producing no shrinkage. In discussion of the results it has been assumed that the light arc is contracting relative to the black; an alternative possibility would be that the black elongates relative

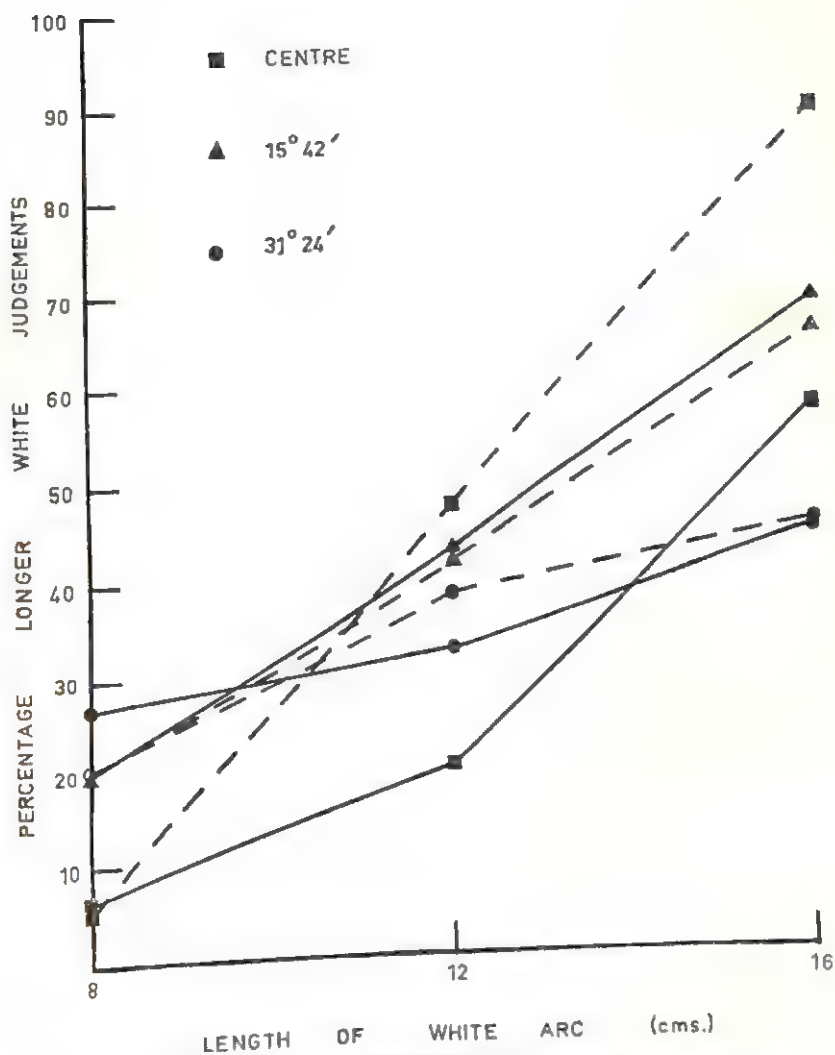


FIGURE 1  
PERCENTAGE OF TIMES THAT WHITE ARCS WERE JUDGED LONGER THAN  
A 12 CM BLACK ARC

Complete lines indicate presence of leading contour, broken lines indicate absence of leading contour.

to the white. However, informal subjective reports suggested that the former was the case in the present experiment.

The effect of the leading contour did not hold for peripheral angles greater than seven degrees 10 minutes (the angle subtended by the path of rotation



in the straight-ahead fixation position), whereas the standard Ansbacher effect occurs at 12 degrees (1) and also has been reported to occur at angles up to 30 degrees (5). Hence the implications of the present result for an understanding of the standard effect are not immediately apparent.

### E. SUMMARY

Twenty Ss were required to estimate the relative lengths of the white and black teeth on a Sherrington disc rotating at 90 rpm. Judgments were obtained with a standard black tooth of 12 cm in length and a white tooth varying from eight to 16 cm in length. When the disc-centre, which subtended a visual angle of seven degrees 10 minutes to the teeth, was fixated under conditions of a leading but no trailing contour, there were significantly fewer white longer judgments than under conditions of a trailing but no leading contour. Apparent shrinkage was occurring in the former condition, but not in the latter. As the peripheral angle of viewing was increased, the presence or absence of a leading contour produced little difference in the relative judgments of the arc lengths.

### REFERENCES

1. ANSBACHER, H. L. Distortion in the perception of real movement. *J. Exper. Psychol.*, 1944, **34**, 1-23.
2. BEKESY, G. VON. Experiments in Hearing. New York: McGraw-Hill, 1960. Pp. 628.
3. MARSHALL, A. J., & STANLEY, G. The apparent length of light and dark arcs seen peripherally in rotary motion. *Aust. J. Psychol.*, 1964, **2**, 120-128.
4. SHERRINGTON, C. S. On reciprocal action in the retina as studied by means of some rotating discs. *J. Physiol.*, 1897, **21**, 33-54.
5. STANLEY, G. A study of some variables influencing the Ansbacher shrinkage effect. *Acta Psychol.*, 1964, **22**, 109-118.
6. ———. Apparent shrinkage of a rotating arc as a function of luminance relations between figure and surround. *Acta Psychol.*, 1966, **25**, 357-364.
7. ———. Apparent brightness of a rotating arc-line as a function of speed of rotation. *Acta Psychol.*, 1967, **26**, 17-21.
8. ———. Light summation and the perceived length of moving lines. *Acta Psychol.*, 1967, **26**, 260-264.

Department of Psychology  
University of New England  
Armidale, New South Wales, 2351  
Australia

## AN INVESTIGATION OF SOME INTRINSIC VARIABLES AFFECTING TEST RESPONSES\*

*Department of Psychology, St. John's College, University of Manitoba*

---

JOHN BOTTRILL

---

### A. INTRODUCTION

Studies have shown that the affective connotation of words is a factor in free-recall learning (2). The form of the test has also been found to evoke different types of response (1). Forced-choice questions are more likely to produce errors—i.e., change—while open-ended questions are relatively free from them (3).

As a follow-up to Vernon's work, two experiments were carried out by the author to measure the effects of the variables, emotionality, type of test (multiple-choice, rating scale, or straight definition), and speed on the dependent variable—i.e., the mean change—as a result of the variables, between the two identical tests in which Ss were asked to define a set of stimulus words.

### B. EXPERIMENT I (PILOT STUDY)

#### 1. Method

Twenty students, male and female, aged 18-25 years, from a social psychology course served as Ss.

Tests A<sub>1</sub> and A<sub>2</sub> were identical tests asking S to define 20 words of varying affect by responding to the question "What is . . . ?" (e.g., peace). Test B was a straight definition-type questionnaire in which a nonemotional statement was made about each of the 20 words, and Ss were asked to indicate their agreement on a five-point scale: e.g., "Peace is a harmony existing between two or more beings or groups of beings." Test C was a multiple-choice questionnaire in which Ss were asked to choose the statement, referring to one of the 20 stimulus words, with which they most agreed. Test C statements were more emotional than those in Test B.

Each S was given the three tests in one of the three sequences—A<sub>1</sub>BA<sub>2</sub>, A<sub>1</sub>CA<sub>2</sub>, or A<sub>1</sub>BCA<sub>2</sub>. The amount of change from Test A<sub>1</sub> to A<sub>2</sub>, with the

---

\* Received in the Editorial Office, Provincetown, Massachusetts, on October 24, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

intervening activity, was judged by the author on a five-point scale ranging from 1 (same) to 5 (completely different).

## 2. Results

Results of the pilot study showed that there was a greater mean change for the sequences  $A_1BA_2$  and  $A_1BCA_2$  ( $\bar{X}_{A_1BA_2} = 2.45$ ,  $s = .62$ ;  $\bar{X}_{A_1BCA_2} = 2.45$ ,  $s = .44$ ) than for  $A_1CA_2$  ( $\bar{X}_{A_1CA_2} = 2.07$ ,  $s = .49$ ). When the confidence limits were estimated for the population mean ( $\mu$ ) in each of the three sets of tests, those of  $\mu_{ABA}$  and  $\mu_{BCA}$  were found to overlap completely ( $2.18 < \mu_{ABA} < 2.72$ ,  $2.36 < \mu_{BCA} < 2.54$ ,  $p < .05$ ), while  $\mu_{ACA}$  was unrelated ( $1.85 < \mu_{ACA} < 2.29$ ,  $p < .05$ ).

It was concluded from the results that either the form, multiple-choice-type questions, or the degree of affect on Test C brought about less change between the A tests than definition-type questions (Test B).

## C. EXPERIMENT II

As a result of the pilot study it was hypothesized that (a) the type of test, multiple-choice or definition; (b) the degree of emotionality or affect of the stimulus word itself—e.g., peace; (c) the affect of the definition of the word; and finally (d) the completion speed would have an effect on the amount of change between Tests  $A_1$  and  $A_2$ . To test these hypotheses, a new list of words and five improved questionnaires were constructed, as described below.

### 1. Method

A group of 96 students, male and female, aged 18-25 years, from two sections of an introductory psychology course served as Ss in this experiment.

In order that the improved questionnaire might be constructed, a group of 14 male and female students, aged 18 to 25 years, was asked to categorize a list of 34 common words as either emotional or nonemotional. The words with the greatest frequency of emotional or nonemotional scores were chosen as the 20 words to be used in the new questionnaire. This resulted in 11 emotional and nine nonemotional words.

Emotional and nonemotional definitions were made up for both the multiple-choice (Test C) and the definition (Test B) tests for all 20 stimulus words in the new Test A. An emotional definition was one designed to arouse an emotional response—e.g., disgust, pleasure, fear—and a nonemotional was one that did not. Instructions on the tests were the same as in Experiment I.

The five new questionnaires were (a) Test A with the new topic words,

(b) a definition-type with emotional definitions ( $B_E$ ), a definition-type with nonemotional definitions ( $B_N$ ), (c) multiple-choice with emotional definitions ( $C_E$ ), and (d) multiple-choice with nonemotional definitions ( $C_N$ ).

The procedure was the same as in Experiment I except that the Ss were randomly divided into five groups instead of three. The first and last tests were again A questionnaires, and the middle test was  $B_E$ ,  $B_N$ ,  $C_E$ ,  $C_N$ , or a maze. (The control group of Ss did the maze as a filler activity between the two  $A_1$  tests, in order that the time might be kept approximately equal for all five testing combinations: i.e.,  $A_1$  maze  $A_2$ ,  $A_1B_EA_2$ ,  $A_1B_NA_2$ ,  $A_1C_NA_2$ , and  $A_1C_EA_2$ .)

The stapled test sequences were randomly distributed among the Ss in each group. They were told that the test was a kind of attitude test and also that each person would have a test different from that of the person beside him. The Ss were not required to give their names. They were instructed not to refer back to the previously completed questionnaires and told that they must answer all questions.

The speed of completion was recorded by the experimenter. Fast Ss were those finished before 20 minutes, Medium before 35 minutes, and Slow after 35 minutes.

## 2. Results

The mean change with the sequence  $A_1$  maze  $A_2$  was 2.1 ( $s = .43$ ), and the mean change of Tests  $A_1B_{E\&N}A_2$  ( $\bar{X} = 2.6$ ) and  $A_1C_{E\&N}A_2$  ( $\bar{X} = 2.55$ ) were significantly different ( $1.9 < \mu_{A_1 \text{ maze } A_2} < 2.3$ ,  $p < .05$ ) but  $A_1B_{E\&N}A_2$  and  $A_1C_{E\&N}A_2$  did not differ significantly from each other ( $F = 1.16$ ,  $df = 1,236$ ,  $p > .05$ ). Thus, the null hypothesis, that multiple-choice and definition-type questions do not differ in their effects on Test  $A_1$  and Test  $A_2$  cannot be rejected, contrary to what was originally expected.

In regard to  $H_2$  it was found that the emotionality or nonemotionality of the stimulus word did not significantly affect the mean change in any of the test sequences ( $F_{A_1 \text{ maze } A_2} = .09$ ,  $F_{A_1B_{E\&N}A_2} = .46$ ,  $F_{A_1C_{E\&N}A_2} = .17$ ,  $p > .05$ ).

A two-way analysis of variance showed that for Test C there was a significant difference ( $F = 4.67$ ,  $df = 1,114$ ,  $p < .01$ ) between the effects on  $A_2$  of the emotional and nonemotional tests; and for Test B, almost a significant difference ( $df = 1,114$ ,  $F = 3.65$ ). There was a significant interaction effect between the emotional Test C and the nonemotional Test B and the speed ( $df = 2,114$ ,  $F = 8.86$ ,  $p < .001$ ), though this did not show up with Test B.

For Tests  $A_1$  maze  $A_2$  and  $A_1C_{E\&N}A_2$ , the speed of completion was statis-

tically significant ( $F_{AA} = 4.61$ ,  $df = 2,48$ ,  $p < .05$ ;  $F_{AC_{E\&N}A} = 5.65$ ,  $df = 2,114$ ,  $p < .01$ ); and for Test  $AB_{E\&N}A$ ,  $F_{ABA} = 2.37$ ,  $df = 2,114$ ,  $p > .05$ . On Test  $A_1$  maze  $A_2$ , comparison of the mean change of the Fast ( $\bar{X} = 2.4$ ,  $s = .51$ ) and Medium ( $\bar{X} = 1.9$ ,  $s = .36$ ) finishers gave  $t = 3.33$  ( $df = 17$ ,  $p < .01$ ) and of the mean changes of the Fast and Slow ( $\bar{X} = 2.1$ ,  $s = .37$ ) finishers gave  $t = 2.0$  ( $df = 17$ ,  $p \approx .05$ ). The relationship was not significantly curvilinear ( $\eta = .42$ ,  $F = .63$ ). Thus speed of completion does affect the amount of change.

#### D. DISCUSSION

It is clear from these results that the kind of intervening activity between Tests  $A_1$  and  $A_2$  does make a difference to the mean amount of change. There was a greater amount of change between the tests when the activity was a multiple-choice or definition-type questionnaire than when it was a maze "filler" activity.

Contrary, however, to the hypothesis, the different types of questionnaire (multiple-choice and definition) did not differ in their effect on the amount of change produced. We can therefore reject the hypothesis that Test B will produce more change than C, thus disproving Anastasi and Vernon, who stated that there is a difference in the effects of multiple-choice and straight definition-type questions. It is possible, however, that this difference would show up in another test and that it was insignificant in this experiment because of inappropriate measurements.

It was found that the degree of emotionality of the stimulus word did not make a significant difference in the amount of change between the tests  $A_1$  and  $A_2$ . It is possible that the random order of the emotional and nonemotional topics on the tests nullified any effect they might have had. If the tests had been given to a different population—for example, emotionally-disturbed children, or perhaps a group of students from a different culture—there might have been more or less change in the results as a direct result of the emotionality of the topics themselves.

The emotionality of the questionnaire, more for the multiple-choice than the definition test, makes a difference in the amount of change. When the test is definition-type, there is less of a difference between the effects of emotional and nonemotional questions than when the test is multiple-choice.

There was a significant interaction effect between the emotionality/non-emotionality of Test C and the completion speed. Slow Ss had a greater mean change on the nonemotional questionnaire, while the fast Ss had a greater



mean change on the emotional questionnaire. The relationship between speed of completion and emotionality is discussed below.

Ss who finish quickly were shown to have a greater amount of change from Test A<sub>1</sub> to A<sub>2</sub> than those who finished slowly, and both the slow and fast finishers changed more between the tests than the Ss who took a medium length of time to finish. The fast-finishers may perhaps be more field dependent and extrovert, and be more interested in just completing the task quickly than trying to answer the questions to the best of their ability: i.e., task-oriented. The slow finishers, alternatively, may take things slowly and calmly and not be affected by either the field or task completion: i.e., ego-oriented. It is also possible that the slow finishers spend so much time trying to remember what answers they gave to the first test, that they do not produce much change.

### E. SUMMARY

The effects of independent variables—e.g., speed of completion, type of filler activity, emotionality of the stimulus words, and degree of affect in their definitions were measured by the amount of change in response shown by psychology students on two administrations of an open-ended questionnaire.

It was found that all the variables except emotionality of the stimulus word had an effect on the second set of responses, for these particular subjects, and that interaction effects need to be taken into account.

### REFERENCES

1. ANASTASI, A. *Psychological Testing*. New York, N.Y.: Macmillan, 1961. P. 512.
2. McNULTY, A., & ISNOR, C. J. E. Affective connotation of words as a factor in free-recall learning. *Brit. J. Psychol.*, 1967, 58(1), 21-57.
3. VERNON, P. E. *The Measurement of Abilities*. London, England: Univ. London Press, 1940. P. 224.

*Department of Psychology*  
*Brandon University*  
*Brandon, Manitoba, Canada*



## S-R AND R-S PAIRED-ASSOCIATE LEARNING AS A FUNCTION OF CONCRETENESS, IMAGERY, SPECIFICITY, AND ASSOCIATION VALUE\*<sup>1</sup>

*Department of Psychology, University of Guelph, Canada*

A. DANIEL YARMEY AND BRIAN J. O'NEILL<sup>2</sup>

### A. INTRODUCTION

Several investigators have demonstrated that conditioned sensory images elicited by paired-associate (PA) nouns act as mediators for recall in PA verbal learning (8, 21). An image common to both members of pairs is said to mediate an association between them and to facilitate recall. Paivio, Yuille, and Madigan (13) have made explicit the implication that the image arousing capacity of words depends on their directness of sensory reference: ratings of noun concreteness (C) and noun imagery (I) were found to correlate .83 in a sample of 925 nouns, suggesting that the two variables define the same underlying dimension. A study by Yuille (23) presented evidence that I is the determining variable responsible for any influence of C in PA learning.<sup>3</sup>

Paivio (9) has hypothesized that conditioned imagery elicited by PA members forms a compound imaginal mediator, thus facilitating learning. Since concrete noun stimuli acting as "conceptual pegs" would serve to reinstate mediating images from which response terms may be decoded as words, C effects would be stronger on the stimulus than on the response side of pairs. A number of studies have found support for this hypothesis [see (9) for the most complete summary of the literature].

Theoretically, a continuum of image arousal capacity of stimuli follows the order, from strongest to weakest: the object itself, pictures, concrete specific nouns (e.g., hammer), concrete general (tool), abstract specific nouns (algebra), abstract general (mathematics), high association value (AV) nonsense syllables, and low AV nonsense syllables. Recall of concrete noun-noun pairs

\* Received in the Editorial Office, Provincetown, Massachusetts, on October 28, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This study was supported by a grant to the senior author from the National Research Council of Canada (APA 288). The authors wish to thank Dick de Ryk, Peggy Klinck, and Anita Stewart for their assistance.

<sup>2</sup> Now at the University of Western Ontario.

<sup>3</sup> Since I and C are so closely related, reference to either one throughout the paper implicitly includes reference to the other.

is relatively easily explained by the conceptual peg theory; it is more difficult to interpret the recall of abstract words and nonsense syllables (NS) at the low end of the imagery continuum. A high rated I stimulus term should facilitate recall of low rated I nouns and even NS provided that the response terms readily elicit word associations which themselves elicit images. In this manner, a compound stimulus-response image could be formed; the stimulus term would reinstate the compound image from which the verbal association and ultimately the response term could be decoded.

All of the PA studies testing the mediating imagery theory have employed the A-B forward direction paradigm. In order to test further the theoretical significance of the imagery hypothesis, R-S learning, in addition, should be examined. The present study was designed to investigate the strength of S-R and R-S bonds with different combinations of stimulus and response, C, I, specificity of nouns, and AV of nonsense syllables.

#### B. EXPERIMENT 1<sup>4</sup>

Backward associative learning, as Ekstrand (4) pointed out, consists of two stages: stimulus learning and B-A associative learning. Concrete nouns, as opposed to abstract nouns, are known to have a differential effect on stimulus learning. Before testing for B-A strength, a correction for the unequal stimulus learning is required. The procedure used by Weiss (18) was employed in the present study. Following acquisition of A-B pairs, half of the Ss were transferred to B-A learning and the remainder to B-A re-paired (B-Ar) learning. In the B-Ar condition the former stimuli and responses of the A-B list are reversed and, in addition, re-paired. Item availability should be equal for both conditions, since the same responses are employed in both paradigms. On the backward association test, the more potent the association, the greater would be the positive transfer in the B-A paradigm and the greater the negative transfer in the B-Ar paradigm. A conversion of scores to B-A/B-Ar ratios would represent R-S associative strengths. The greater the effect of the independent variable, the smaller would be the ratio in percentage scores.

Study one examined the effect of learning concrete-concrete (C-C), concrete-abstract (C-A), abstract-concrete (A-C), and abstract-abstract (A-A) noun PAs. It was hypothesized on the basis of previous research (8) and the conceptual peg theory that Ss would have the least difficulty in learning the C-C pairs and increasing difficulty for C-A, A-C, and A-A pairs. For R-S

<sup>4</sup> Experiment I is based on the M.A. thesis of the second author conducted under the supervision of the first author.

learning, C-C pairs should be the easiest and A-A pairs the most difficult to acquire because of the greater availability in recall of concrete noun pairs over abstract pairs. If increasing stimulus availability (via high I, concrete nouns) during S-R learning is manifested in superior R-S learning, the test for R-S strength should be greater for the re-located PAs, A-C than C-A, since the high I noun responses are to be recalled. However, if stimulus imagery of S-R learning does not influence R-S associative learning, the effects being a function of stimulus imagery during the R-S recall test, the reversed pairs C-A should be better recalled than reversed pairs A-C.

### 1. Method

*a. Subjects.* One hundred and twenty undergraduates, all inexperienced in verbal learning experiments, served as Ss, each S being assigned to one of four groups according to his order of appearance at the laboratory. An additional seven Ss were excluded from the experiment; three because of equipment failure, four because of language difficulties. When an S was dropped from the experiment, the next S to appear at the laboratory was put into the same experimental condition.

*b. Materials.* The concrete nouns (garden, avenue, factory, beast, hotel, corner, boulder, journal, mother, slave, dress, furniture, lad, nail, clothing, string, flesh, student, village, and engine) and abstract nouns (theory, law, knowledge, amount, event, quality, justice, position, soul, honor, custom, chance, style, history, moral, advice, attitude, freedom, duty, and effort) employed were chosen from the Paivio, Yuille, and Madigan (13) norms on which I, C, and meaningfulness (*m*) scores were available. All of the concrete nouns exceeded the abstract in rated I. The mean C and I scores of the concrete nouns are 6.71 and 6.27, respectively; and of abstract nouns, 2.39 and 3.14. The mean *m* of the concrete nouns is 5.98; of the abstract nouns, 5.91. The mean word length for both concrete and abstract nouns is six letters. All of the nouns are of A or AA word frequency (15).

The PA lists, each consisting of 10 pairs, were constructed so that each list consisted of C-C, C-A, A-C, or A-A noun pairs. The A-C pairs were the C-A pairs "flipped over." The pairing of individual items was random with the proviso that no obvious mnemonic connection existed between the members of any pair.

*c. Procedure.* Standard instructions for the anticipation method for PA learning were given to each S by tape-recordings. The lists were presented at a 2:2 second-rate on a memory drum. There was a 4-second intertrial interval for both S-R and R-S learning and a 2-minute between-stage interval. Subjects



were given one minute to free recall the PAs, and one minute was taken to change the tapes on the memory drum. Four randomizations of the position of pairs in each list were used to minimize serial learning. The criterion of learning for both stages was two consecutive errorless trials. The first trial of R-S learning for the experimental groups constituted an aided recall (AR) trial: i.e., Ss were told to anticipate the new responses when given the previous responses from S-R learning as stimuli. The Ss assigned to R-Sr learning were not tested for AR. Subjects were not told during the instructions that they would be tested for free recall, AR, or backward learning.

## 2. Results and Discussion

a. *S-R learning.* The total errors to criterion for each S were determined. A comparison of total error scores between Ss assigned to R-S and R-Sr conditions indicated no significant differences for S-R learning for each of C-C ( $t < 1$ ), C-A ( $t_{28} = 1.82$ ) A-C ( $t_{28} = 1.86$ ) and A-A ( $t < 1$ ). Learning scores for Ss given the R-Sr conditions were considered to be of interest only for the backward learning stage and were excluded from further S-R analysis. The means and standard deviations (SDs) of error rate, respectively, were C-C, 38.13, 24.94; C-A, 32.47, 19.87; A-C, 53.13, 35.89; A-A, 118.00, 72.02 ( $n = 60$ ). These data were analyzed by an analysis of variance with stimulus concreteness and response concreteness treated as between-Ss factors. Significant effects were attributable to stimulus concreteness ( $F = 17.05$ ,  $df = 1/56$ ,  $p < .001$ ), response concreteness ( $F = 5.49$ ,  $df = 1/56$ ,  $p < .05$ ), and their interaction ( $F = 8.03$ ,  $df = 1/56$ ,  $p < .01$ ). An *a posteriori* test of the difference among means (3), showed that the A-A treatment differed from each of the other treatments, C-C, C-A, and A-C ( $p < .05$ ). No other differences among means were statistically significant.

These findings are consistent with previous studies, which show that concrete nouns are more facilitatory for learning than abstract nouns. However, unlike some other reports, C was facilitatory on the response side of pairs, as well as on the stimulus side. Paivio (8), using the study-recall method, found similar results, but argued that the findings still supported the conceptual peg hypothesis. Since the variance on the stimulus side of the pairs was eight times greater than on the response side, stimulus C and I were actually greater aids for learning than response C and I. In the present study, the variance attributable to C on the stimulus side of the pairs was three times greater than on the response side.

b. *Free recall.* The means and SDs of correctly recalled pairs for experimental conditions, respectively, were C-C, 7.73, 1.39; C-A, 6.73, 2.67; A-C,

7.67, 2.36; and A-A, 6.67, 1.81. An analysis of variance for these scores yielded no significant differences, which suggests that the pairs were equally available for all conditions following S-R learning.

*c. R-S learning.* Table 1 presents the mean scores for AR and total errors to criterion for R-S and R-Sr learning. The table shows the PAs in their reversed backward order so that former response items now serve as stimuli. The first analysis was confined to R-S learning, excluding R-Sr scores. On both AR and total errors to criterion, the only significant effect found was for stimulus concreteness ( $F = 11.20$ ,  $df = 1/56$ ,  $p < .01$ , for AR;  $F = 13.04$ ,  $df = 1/56$ ,  $p < .001$ , for total errors). These findings give strong support to a relocated "conceptual peg" effect acting on and facilitating R-S recall and learning.

TABLE 1  
MEAN SCORES FOR AIDED RECALL (AR) AND TOTAL ERRORS (TE) TO CRITERION FOR  
R-S AND R-Sr LEARNING USING CONCRETE AND ABSTRACT NOUNS

Stimulus	Measure	Pairs	Response			
			Concrete		Abstract	
			$\bar{X}$	$SD$	$\bar{X}$	$SD$
Concrete	AR		8.27	1.61	8.33	1.45
		R-S	4.07	4.74	4.20	3.47
	TE	R-Sr	71.87	29.73	73.40	42.76
Abstract	AR		7.20	2.00	6.00	2.39
		R-S	14.07	21.10	28.61	20.16
	TE	R-Sr	102.00	24.04	129.00	52.02

The next analysis was a comparison of the total error scores for R-S and R-Sr learning. A  $2 \times 2 \times 2$  analysis of variance with stimulus concreteness, response concreteness, and experimental-interference conditions treated as between-Ss variables yielded significant main effects for stimulus concreteness ( $F = 27.95$ ,  $df = 1/112$ ,  $p < .001$ ) and experimental-interference conditions ( $F = 201.23$ ,  $df = 1/112$ ,  $p < .001$ ), but not for response concreteness ( $F = 3.44$ ). The important effect in terms of the Weiss (18) paradigm was the stimulus concreteness  $\times$  experimental-interference interaction ( $F = 4.88$ ,  $df = 1/112$ ,  $p < .05$ ). These results indicated that pairs involving concrete stimuli and abstract stimuli differed in their susceptibility to positive and negative transfer conditions. The R-S/R-Sr error ratio percentage scores for the conditions C-C, C-A, A-C, and A-A of 6, 6, 15, and 22 respectively, indicate that transfer effects were stronger for concrete stimulus pairs than for abstract stimulus pairs. It is clear that relocated peg effects and not transfer factors from differential stimulus imagery during S-R learning, as indicated

by the superior performance of C-A over A-C learning, determined the strength and the direction of backward associations. Thus, the hypotheses for Experiment I were fully supported for R-S learning, while only partially supported for S-R learning.

### C. EXPERIMENT II<sup>5</sup>

The second experiment was designed to examine S-R and R-S learning with the use of concrete nouns and abstract nouns as stimuli or responses with NS—AV 50-60 per cent (7)—as associates. On the basis of the conceptual peg hypothesis and the findings of Paivio and Madigan (11), Yarmey (20), and Yarmey and Paivio (22), it was predicted that (a) S-R learning of concrete noun (CN)-NS pairs would be superior to learning of abstract noun (AN)-NS pairs; (b) no difference in acquisition rate was expected between NS-CN and NS-AN pairs; (c) concreteness would aid R-S associative learning; and (d) facilitated backward recall and learning would occur, depending on the locus of the CNs. Thus, AR and learning rate for relocated R-S pairs should be promoted when CNs, as opposed to ANs, act as stimuli for the NS responses.

#### 1. Method

a. *Subjects.* One hundred and twelve undergraduates served as Ss and were assigned alternately in order of appearance at the laboratory to one of four S-R conditions: CN-NS, AN-NS, NS-CN, and NS-AN. Five additional Ss were excluded from the experiment, two for failure to understand the instructions, two for apparatus difficulties, and one for health reasons.

b. *Materials.* Ten concrete nouns, abstract nouns, and NS were used to form four PA lists. The NS-CN and NS-AN PAs were the identical but "turned over" terms used to form the CN-NS and AN-NS pairs. The concrete nouns (breast, poet, blood, officer, dress, automobile, author, factory, circle, and professor) and abstract nouns (duty, effort, theory, justice, method, thought, chance, position, style, and virtue) were chosen from the Paivio, Yuille, and Madigan (13) norms. All words are highly familiar—(A or AA, Thorndike-Lorge (15))—and were equated for *m* (concrete  $\bar{x} = 5.65$ , abstract  $\bar{x} = 5.69$ ). The mean C and I scores of the concrete nouns are 6.57 and 6.17, respectively; and of abstract nouns, 2.16 and 3.07. The NS (BEJ, GUK, DAQ, FOH, PUV, HYW, QOG, JEX, MAF, NIS) employed have AVs of 50-61 per cent (7). The same NS served as associates for both concrete and abstract nouns.

<sup>5</sup> Results of Experiment II were presented by the first author at the meeting of the Canadian Psychological Association, Calgary, Alberta, June 1968.

*c. Procedure.* The only difference in procedure between experiments I and II was the criterion for learning. For both S-R and R-S stages, the PAs were learned to a criterion of one perfect trial or 40 trials, whichever came first. Subjects were required to pronounce both the words and NS anticipations.

## 2. Results and Discussion

*a. S-R learning.* An analysis of total errors to criterion for Ss assigned to R-S and R-Sr treatments yielded no significant differences for NS-CN ( $t < 1$ ), CN-NS ( $t < 1$ ), and NS-AN ( $t < 1$ ). However, Ss assigned to the interference condition for learning of AN-NS pairs made significantly fewer errors than those in the experimental condition (the means were 92.21 and 126.43, respectively,  $t_{26} = 2.41$ ,  $p < .05$ ). Since Ss were assigned to conditions randomly, by order of appearance at the laboratory, the difference in learning ability between Ss in the R-S and R-Sr treatments for the AN-NS list can only be accounted for as chance. This discrepancy must limit the significance of any differences in which concreteness and R-S and R-Sr learning are compared. Learning scores for Ss assigned to the R-Sr treatment were not included in the analysis of S-R performance.

The means and *SDs* of total errors to criterion for each experimental group, respectively, were CN-NS, 55.00, 22.26; AN-NS, 126.43, 36.23; NS-CN, 90.57, 61.13; and NS-AN, 112.50, 40.38 ( $n = 56$ ). An analysis of variance, with concreteness and stimulus-response position treated as between-Ss factors, yielded significant main effects only for concreteness ( $F = 15.79$ ,  $df = 1/52$ ,  $p < .01$ ). The interaction of position and concreteness was also significant ( $F = 4.44$ ,  $df = 1/52$ ,  $p < .05$ ). A comparison of means by  $t$  tests showed that Ss learning CN-NS made fewer errors than Ss learning AN-NS ( $t = 6.06$ ,  $df = 26$ ,  $p < .001$ ). The error scores between NS-CN and NS-AN yielded no significant difference ( $t = 1.08$ ).

The superiority in PA learning of CNs over ANs as stimuli, but not as responses for NS associates, supports hypotheses one and two. This experiment both sustains and extends the hypothesis that stimulus term imagery would facilitate recall in PA learning when response terms are not real words.

Paivio and Madigan (11) found that high AV syllables—93 per cent (5)—were more easily recalled as associates to CN stimuli than low AV syllables—17 per cent or less (19). Yarmey and Paivio (22) found similar results employing NS of 70-80 per cent AV (7). The present study used NS of an intermediate AV range, 50-61 per cent (7). It may be concluded that if the response items of the PAs are of at least intermediate AV, they may be easily transformed into words which themselves may elicit images. The compound



stimulus-response image could then be formed and thereby facilitate response recall.

*b. Free recall.* The means and *SDs* of correctly recalled pairs for experimental conditions, respectively, were CN-NS, 8.36, 1.63; AN-NS, 6.36, 2.77; NS-CN, 7.28, 1.71; and NS-AN, 6.21, 0.86. An analysis of variance indicated that the only significant effect was for concreteness ( $F = 8.78$ ,  $df = 1/52$ ,  $p < .01$ ), indicating that CNs and NS associates, regardless of their stimulus-response position, were more available than ANs and NS associates following S-R learning.

*c. R-S learning.* Table 2 presents the mean scores for AR and total errors to criterion for R-S and R-Sr learning. An analysis of variance performed on the AR scores supported hypothesis three; concrete pairs were better recalled than abstract pairs ( $F = 4.32$ ,  $df = 1/52$ ,  $p < .05$ ). Backward recall of nouns was superior to backward recall of nonsense syllables ( $F = 10.39$ ,  $df = 1/52$ ,  $p < .05$ ). The interaction of concreteness  $\times$  stimulus-response position fell just short of statistical significance ( $F = 3.74$ ,  $df = 1/52$ ,  $p < .10$ ). Since the interaction was of interest, it was decided to compare the means by *t* tests. The AR of NS responses for concrete noun stimuli was significantly superior to recall of NS responses with abstract noun stimuli ( $t = 3.40$ ,  $df = 26$ ,  $p < .01$ ). No significant difference was found between NS-CN and NS-AN PAs ( $t < 1$ ). The measure of AR thus lends only partial support to hypothesis four.

TABLE 2  
MEAN SCORES FOR AIDED RECALL (AR) AND TOTAL ERRORS (TE) TO CRITERION FOR  
R-S AND R-Sr LEARNING USING NOUNS (N) AND NONSENSE SYLLABLES (NS)

Locus	Measure	Pairs	Lists			
			Concrete		Abstract	
			$\bar{X}$	<i>SD</i>	$\bar{X}$	<i>SD</i>
NS-N	AR	R-S	8.21	2.30	8.14	1.77
		R-Sr	2.71	3.10	10.07	11.34
	TE	R-S	62.14	28.25	75.14	34.99
		R-Sr	7.57	1.64	5.57	1.35
N-NS	AR	R-S	8.07	11.63	31.00	16.71
	TE	R-Sr	121.00	61.15	195.28	82.16

Hypothesis three was also supported for R-S learning. An analysis of variance of total errors to criterion yielded significant main effects attributable to concreteness ( $F = 21.58$ ,  $df = 1/52$ ,  $p < .001$ ), noun-nonsense syllable position ( $F = 16.25$ ,  $df = 1/52$ ,  $p < .001$ ), and the interaction of concreteness and position ( $F = 5.70$ ,  $df = 1/52$ ,  $p < .05$ ). Concrete nouns were clearly



superior to abstract nouns as stimuli ( $t = 4.06$ ,  $df = 26$ ,  $p < .001$ ) and as responses, although at a lesser degree of significance ( $t = 2.26$ ,  $df = 26$ ,  $p < .05$ ). These measures clearly indicate that C facilitated R-S learning. Learning of R-S pairs also indicated that the relocated stimulus imagery or "peg effect" differentially facilitated performance. However, since CNs were more available than ANs following S-R acquisition as measured by free recall, these results must be accepted with reservation.

The differential strength of backward learning was also examined comparing R-S and R-Sr PA learning ( $n = 112$ ). Two separate analyses of variance were conducted in which the variables, concreteness-nonsense syllables and experimental-interference conditions, were treated as between-Ss variables. The only significant effect found for the analysis of total error scores of NS-CN and NS-AN backward learning was attributable to the experimental-interference factor ( $F = 93.25$ ,  $df = 1/52$ ,  $p < .001$ ). The R-S/R-Sr error ratios for NS-CN and NS-AN of four per cent and 13 per cent were considered to be equivalent, since the concreteness-syllable  $\times$  experimental-interference interaction was not significant. The analysis of total error scores for CN-NS and AN-NS backward learning indicated that learning with concrete nouns as stimuli was superior to learning with abstract nouns as stimuli ( $F = 11.10$ ,  $df = 1/52$ ,  $p < .01$ ). The R-S pairs were easier to learn than R-Sr pairs ( $F = 92.05$ ,  $df = 1/52$ ,  $p < .01$ ), but the interaction of concreteness  $\times$  R-S/R-Sr was not significant ( $F = 3.23$ ). The R-S/R-Sr error ratio scores for CN-NS and AN-NS of seven per cent and 16 per cent, therefore, must be considered as equivalent. Although these results may be interpreted as suggesting that stimulus imagery of S-R learning has no influence on R-S associative strength, another interpretation is more plausible. It is likely that high I nouns affect the interference of re-pairing. Consequently, there would be no differential interference for CNs and ANs and no significant difference in the interaction of concreteness  $\times$  R-S/R-Sr conditions. It may be concluded that CN stimuli are superior facilitators for learning than AN stimuli in both R-S and R-Sr learning.

#### D. EXPERIMENT III

Paivio and Madigan (11) interpreted the inferior recall of low AV syllable responses as opposed to high AV syllables as a failure by Ss to encode the NS into words, thus limiting any possible facilitation for learning by stimulus mediating imagery. The present study attempted to examine the influence of low AV NS and C on S-R and R-S learning. No differential effect was expected for C as stimuli on either S-R or R-S learning.

### 1. Method

The experimental design was similar to Experiment II with the following exceptions. Forty-eight undergraduates served as Ss. The R-Sr learning condition was not included. The criterion of learning for both S-R and R-S stages was 8/10 correctly anticipated pairs or 40 trials, whichever came first.

The concrete nouns and abstract nouns used in Experiment II were employed. The 10 CCC trigrams (XIDL, WGP, HZW, SGJ, CQZ, NCF, GKQ, KXR, FPJ, and BWF) employed have AVs of 17 per cent or less (19). Subjects were required to spell out the nonsense syllable anticipations and pronounce the words.

### 2. Results and Discussion

*a. S-R learning.* The means and *SDs* of total errors to criterion for the four groups, respectively, were CN-NS, 212.83, 61.23; AN-NS, 242.25, 63.50; NS-CN, 113.08, 66.32; and NS-AN, 183.25, 68.89. Concrete nouns were again found to be easier to learn than abstract ( $F = 6.44$ ,  $df = 1/44$ ,  $p < .05$ ). An  $F(1,44)$  of 16.83, significant at the .001 level of confidence, confirmed the potent influence of high *m* (trigrams *vs.* nouns) on the response side of pairs (17). The interaction of concreteness  $\times$  stimulus-response position was not significant, confirming the expectation that there would be no difference in learning between concrete and abstract nouns as stimuli with low *m* response items. These findings are in agreement with Paivio and Madigan (11).

*b. Free recall.* The mean and *SD* of correctly recalled pairs for each group, respectively, were CN-NS, 6.08, 1.71; AN-NS, 5.25, 1.53; NS-CN, 2.17, 1.21; and NS-AN, 2.67, 2.28. The only significant effect found was attributable to the locus of nouns on the stimulus side of the PAs ( $F = 38.86$ ,  $df = 1/44$ ,  $p < .001$ ). No differential effect existed for concreteness-abstractness. The use of low AV trigrams as stimuli probably resulted in stimulus fractionation (6). Consequently, when Ss were asked to free recall the pairs, they were unable, due to having learned only the first letter of the trigram stimuli.

*c. R-S learning.* The means and *SDs* for AR for each list, respectively, were CN-NS, 6.58, 1.55; AN-NS, 5.92, 1.98; NS-CN, 1.58, 1.32; and NS-AN, 2.08, 1.70. The means and *SDs* for total errors were CN-NS, 118.17, 57.48; AN-NS, 119.33, 86.51; NS-CN, 14.50, 14.50; and NS-AN, 18.42, 22.85, respectively. For both AR and total errors to criterion, locus of the nouns was the only significant effect ( $F = 78.22$ ,  $df = 1/44$ ,  $p < .001$  for AR;  $F = 39.96$ ,  $df = 1/44$ ,  $p < .001$  for errors).

The results for Experiment III support the hypothesis that stimulus *c* does not differentially influence either S-R or R-S learning when low *m* materials are employed as response associates.

#### E. EXPERIMENT IV

The effects of specificity-generality of concrete and abstract nouns on S-R and R-S learning were investigated. Specificity of nouns refers to particulars, while generality applies to whole classes of instances. Specific, as opposed to general concrete, nouns have been found to facilitate PA learning, the effects being greater on the stimulus than on the response side of pairs (12). In another study of PA learning, Paivio (10) employed a complete factorial design in which abstractness-concreteness and generality-specificity of the nouns were varied on either side of the PA lists. The findings supported the conceptual peg hypothesis. Although response specificity yielded significant results, both stimulus concreteness and stimulus specificity showed much larger facilitating effects for learning.

The present study investigated the effects of learning concrete specific-concrete general (CS-CG), concrete general-concrete specific (CG-CS), abstract specific-abstract general (AS-AG), and abstract general-abstract specific (AG-AS) paired-nouns. It was expected that learning the above S-R pairs by the anticipation method would replicate the findings of Paivio (10). For AR and R-S learning it was predicted, from the conceptual peg hypothesis, that reversal of S-R PAs, CG-CS to CS-CG, would result in greatest learning. It was reasoned that the stimulus side of the pairs would have the highest *I* scores and, therefore, should most easily elicit mediating imagery and response retrieval. Accordingly, AR and learning should decrease with R-S pairs in the following order: CS-CG, CG-CS, AS-AG, and AG-AS.

##### 1. *Method*

The experimental design differed slightly from the preceding experiments. Each of the 48 Ss was given auditory presentation of one of the PA lists by tape-recordings. The nouns were presented at a 2:2 second rate, and the anticipation method was used. The criterion of learning for both S-R and R-S stages was 7/8 correctly anticipated pairs or 40 trials, whichever came first.

The concrete, general, and specific word pairs employed were money and dollar, animal and frog, furniture and table, officer and policeman, beverage and wine, clothing and vest, vessel and steamer, vegetable and potato. Abstract, general, and specific word pairs were knowledge and fact, sensation and warmth, quality and strength, humor and joke, time and hour, justice and

law, season and winter, crime and robbery. According to the Paivio, Yuille, and Madigan (13) norms the mean *C*, *I*, and *m* values, respectively, for the four classes of words are CS, 6.86, 6.55, and 6.82; CG, 6.61, 6.09, and 6.76; AS, 3.84, 4.46, and 6.13; and AG, 2.63, 3.92, and 6.38.

## 2. Results and Discussion

*a. S-R learning.* The means and *SDs* for total errors to criterion, respectively, for each list were CS-CG, 1.25 and 0.92; CG-CS, 1.58 and 1.66; AS-AG, 3.97; and AG-AS, 3.83 and 4.14. A simple one-way classification of analysis of variance yielded an  $F(3,44)$  of 3.37, which is significant at the .05 level of confidence. An *a posteriori* test of the difference among means (3) showed that the CS-CG group differed from each of AG-AS and AS-AG ( $p < .05$ ). No other differences among means were statistically significant.

These results indicated that specificity and concreteness of nouns contributed significantly to learning of PAs. Although the differences among means were not as great as found by Paivio (10), the order of means from most superior to poorest learning—CS-CG, CG-CS, AG-AS, and AS-AG—adds general support to the mediating imagery hypothesis.

*b. Free recall, AR and R-S learning.* The means and *SDs* for free recall scores were CS-CG, 6.33 and 0.34; CG-CS, 6.42 and 0.76; AS-AG, 6.00 and 1.08; AG-AS, 5.75 and 0.92, respectively. No significant differences among groups were found ( $F = 1.32$ ). The analysis of AR scores revealed no reliable differences among group means: CS-CG, 7.83; CG-CS, 7.25; AS-AG, 7.33; AG-AS, 7.33 ( $F = 1.49$ ). The analysis of variance of R-S scores also failed to yield significant differences among group means: CS-CG, 0.17; CG-CS, 0.75; AS-AG, 1.42; AG-AS, 1.33 ( $F = 1.12$ ). It is possible that specificity-generality of concrete and abstract nouns have no differential effect on R-S learning. It is also possible, however, that the relatively few PAs used in the present study did not permit a fair assessment of the effects of concreteness and specificity on R-S learning.

## F. GENERAL DISCUSSION

The present experiments indicated that *C* and *I* of stimulus materials affected both forward and backward associative learning. With the exception of Experiment IV, on specificity and concreteness of paired-nouns, the findings supported the conceptual peg hypothesis for both S-R and R-S learning. The results, as expected from the mediating imagery hypothesis, were strongest for learning of noun-noun PAs and weakest for learning of noun-low AV trigrams. However in Experiment I, the evidence from S-R learning regarding the



conceptual peg hypothesis is equivocal. It was predicted that the order of learning superiority, from strongest to weakest, would be C-C, C-A, A-C, and A-A, stimulus I being the most effective variable. Since the results indicated that noun C was an effective variable irrespective of stimulus-response locus and that response C produced a significant difference between A-C and A-A learning scores, the conceptual peg interpretation must be qualified. An examination of the procedure used in noun-noun PA studies which have found support for the conceptual peg hypothesis (e.g., 14) suggests a possible explanation for the present findings. The anticipation method with a 2:2-second presentation rate employed in this study contrasts with the study-recall procedure of Paivio in which the materials were presented at a 2:8-second rate. Since the recall of imaginably mediated associates is considered to involve multiple stages of coding from words to images and back to words, it may be assumed that the shorter the recall interval, the less effective will be the imaginal mediation-recall process.

It should be noted that not all investigators interpret the superiority of learning concrete nouns over abstract nouns in terms of imaginal mediation. Dominowski and Gadlin (2) argue that differences in learning between concrete and abstract noun pairs should be interpreted in terms of stimulus differentiation which results in less intralist interference for high I materials. In the present experiments, the investigators equated concrete and abstract nouns on *m*, frequency, and word length. In addition, homogeneous lists rather than mixed-lists were employed. All of the above controls would tend to negate a stimulus differentiation hypothesis. As a consequence, the present results were interpreted as supporting the conceptual peg hypothesis.

In accord with Underwood (16) it has been stated that degree of learning, and therefore associative bond strength, depends on the rate of PA acquisition. For Experiment I it follows that C-C, C-A, and A-C forward associative bonds were equal and were stronger than A-A associative bonds. Since forward and backward associative strengths are considered to correlate (4) the order of R-S learning superiority should have corresponded to the S-R order of learning proficiency. In effect, the results indicated that the order of proficiency for R-S experimental groups was determined by conceptual peg effects, learning being facilitated by the relocated concrete stimulus "pegs." This finding in favor of the extension of the conceptual peg hypothesis to backward learning is unconfounded by any other significant effects. The dominance of the backward peg effect is evident in the analysis of R-S/R-Sr error ratio scores and of aided recall data in particular, since no previous experience of pairs in the R-S order had preceded the AR test.



These findings are of relevance to the associative symmetry hypothesis (1). This hypothesis states that, subject to the equal availability of stimulus and response items for backward recall, an association formed in the S-R direction is equally strong in the R-S direction. Even though no attempt was specifically made to test the associative symmetry hypothesis, it is apparent that concrete nouns and abstract nouns contribute differentially to both S-R and R-S bond strength as a function of the locus of the nouns.

### G. SUMMARY

Effects of concreteness (C), imagery (I), and specificity of nouns and association value (AV) of nonsense syllables, on S-R and R-S learning, were examined in four experiments. S-R and R-S paired-associate learning were highly facilitated by stimulus C and I with paired-nouns and nouns-nonsense syllables (7)—50-61 per cent AV—as stimulus materials. No differential effects were found for C and I as stimuli with low AV syllable responses (19)—17 per cent or less—on either S-R or R-S learning. Stimulus specificity of concrete nouns facilitated only S-R learning. The results were interpreted as supporting a theory in which high imagery stimulus items act as effective “conceptual pegs” for their associates for both S-R and R-S associative learning.

### REFERENCES

1. ASCH, S. E., & EBENHOLTZ, S. M. The principle of associative symmetry. *Proc. Amer. Philos. Soc.*, 1962, **106**, 135-163.
2. DOMINOWSKI, R. L., & GADLIN, H. Imagery and paired-associate learning. *Canad. J. Psychol.*, 1968, **22**, 336-348.
3. EDWARDS, A. L. *Experimental Design in Psychological Research* (3rd ed.). New York: Holt, Rinehart, & Winston, 1968.
4. EKSTRAND, B. R. Backward associations. *Psychol. Bull.*, 1966, **65**, 50-64.
5. GLAZE, J. A. The association value of nonsense syllables. *J. Genet. Psychol.*, 1928, **35**, 255-269.
6. JENKINS, J. J. Stimulus “fractionation” in paired-associate learning. *Psychol. Rep.*, 1963, **13**, 409-410.
7. KRUEGER, W. C. F. The relative difficulty of nonsense syllables. *J. Exper. Psychol.*, 1934, **17**, 145-153.
8. PAIVIO, A. Abstractness, imagery, and meaningfulness in paired-associate learning. *J. Verb. Learn. & Verb. Behav.*, 1965, **4**, 32-38.
9. ———. Mental imagery in associative learning and memory. *Psychol. Rev.*, 1969, in press.
10. ———. Paired-associate learning and free recall of nouns as a function of concreteness, specificity, imagery, and meaningfulness. *Psychol. Rep.*, 1967, **20**, 239-245.
11. PAIVIO, A., & MADIGAN, S. A. Imagery and association value in paired-associate learning. *J. Exper. Psychol.*, 1968, **76**, 35-39.
12. PAIVIO, A., & OLVER, M. Denotative-generality, imagery, and meaningfulness in paired-associate learning of nouns. *Psychon. Sci.*, 1964, **1**, 183-184.

13. PAIVIO, A., YUILLE, J. C., & MADIGAN, S. A. Concreteness, imagery, and meaningfulness values for 925 nouns. *J. Exper. Psychol.* (Monog. Suppl.), 1968, 76(No. 1), Part 2.
14. PAIVIO, A., YUILLE, J. C., & SMYTHE, P. C. Stimulus and response abstractness, imagery, and meaningfulness, and reported mediators in paired-associate learning. *Canad. J. Psychol.*, 1966, 20, 362-377.
15. THORNDIKE, E. L., & LORGE, I. The Teacher's Word Book of 30,000 Words. New York: Bur. of Publ., Teachers Coll., Columbia Univ., 1944.
16. UNDERWOOD, B. J. Degree of learning and the measurement of forgetting. *J. Verb. Learn. & Verb. Behav.*, 1964, 3, 112-129.
17. UNDERWOOD, B. J., & SCHULZ, R. W. Meaningfulness and Verbal Learning. Chicago, Ill.: Lippincott, 1960.
18. WEISS, E. Stimulus availability and backward association: Meaning, frequency, and associative bond strength. *Psychol. Rep.*, 1965, 16, 427-437.
19. WITMER, L. R. The association value of three-place consonant syllables. *J. Genet. Psychol.*, 1935, 47, 337-360.
20. YARMEY, A. D. Concreteness, imagery, and temporal factors in paired-associate learning. *Psychon. Sci.*, 1967, 8, 417-418.
21. YARMEY, A. D., & CSAPO, K. G. Imaginal and verbal mediation instructions and stimulus attributes in paired-associate learning. *Psychol. Rec.*, 1968, 18, 191-199.
22. YARMEY, A. D., & PAIVIO, A. Further evidence on the effects of word abstractness and meaningfulness in paired-associate learning. *Psychon. Sci.*, 1965, 2, 307-308.
23. YUILLE, J. C. Concreteness without imagery in PA learning. *Psychon. Sci.*, 1968, 11, 55-56.

*Department of Psychology*  
*University of Guelph*  
*Guelph, Ontario, Canada*



## VISUAL CONFUSION MATRICES: FACT OR ARTIFACT?\*

*Behavioral Research Laboratory, U.S. Army Human Engineering Laboratories,  
Aberdeen Proving Ground, Maryland; and Princeton University*

---

DENNIS F. FISHER, RICHARD A. MONTY, AND SAM GLUCKSBERG

---

### A. INTRODUCTION

In recent years, numerous investigators of short-term memory and related phenomena have made considerable use of letters of the alphabet as stimulus items in a variety of types of studies. Many of these studies, especially those concerned with such factors as visual or auditory confusions in memory, make various assumptions concerning the probability with which perceptual confusions should occur between given letters. These assumptions are usually based on (a) intuitive feelings or *a priori* definitions, such as the deduction that "I" is more likely to be confused with "T" than with "O" because "I and T look alike," whereas "I and O look different" (1, 2); (b) confusion matrices, such as those of Gibson, Osser, Schiff, and Smith (3)—which was a by-product of a study dealing with the development of grapheme discrimination in children (6)—and of Hodge (5)—which was similarly a by-product of a study dealing with the legibility of a uniform-strokewidth alphabet (4); or (c) preliminary studies on a segment of the alphabet of particular interest to the investigation at hand (7).

Each of these approaches involves certain problems. The dangers in the intuitive approach are obvious. On the other hand, the utilization of such matrices as those generated by Gibson or Hodge relies on the implicit assumption that the pattern of confusions between letters of the alphabet is somewhat independent of the method of stimulus presentation employed and of the difficulty of the task. Finally, the preliminary study approach is wasteful because it requires each individual investigator to collect his own confusion data.

The major purpose of the present experiment was to determine if there is any evidence for the common assumption that there exists a basic "pattern of confusions" between upper case letters of the alphabet. The duration of stimulus exposure was manipulated to generate two confusion matrices. Further, the data were collected in a fashion that made direct comparison possible with the data previously generated by Hodge (5) and by Pew and Gardner

---

\* Received in the Editorial Office, Provincetown, Massachusetts, on November 4, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

(8). It was hoped that while the total number of confusions would vary between matrices, a pattern of confusions would emerge that would be independent of the exposure technique employed (tachistoscopic *versus* free viewing) and of overall error rate (manipulated by varying stimulus exposure duration).

## B. METHOD

### 1. Subjects

Fifty U.S. Army enlisted men between the ages of 18 and 25, each with at least 20/20 vision (uncorrected), served as Ss. All Ss spoke and read English as their native language.

### 2. Apparatus

Black Chart-Pak "deca dry" upper case letters, Futura medium font, 36 point (approximately 3/8 inches high) were individually exposed through channel A of a Gerbrands two-channel tachistoscope. Each letter was mounted in the center of a 7-3/4-inch square white card. The intensity of the stimulus field was .0024 mL. A white fixation cross composed of bisecting 5/8-inch lines, 1/16 inch thick, on a black background was presented through channel B. The intensity of the cross was .001 mL.

### 3. Procedure

Ss were dark-adapted for 25 minutes. The experimental room was darkened except for a small red pilot light.

Ss were instructed in the operation of the tachistoscope and then received 18 practice trials with the numerals one through nine. Each numeral was presented once with an exposure time of 400 msec and once with an exposure time of 200 msec. Each stimulus presentation was preceded by a one-second exposure of the fixation cross. Ss who failed to identify properly a majority of the numerals were disqualified from further testing. Upon successful completion of the practice trials, each S was given two blocks of 52 trials each (two exposures of each letter) utilizing the upper case letters described above. One block of trials was presented with an exposure time of 200 msec, the other with an exposure time of 400 msec. Blocks (exposure times) were counterbalanced across Ss. The sequence in which the individual letters were presented was determined at random, independently for each S. Each letter of the alphabet was presented a total of 100 times (across Ss) at each exposure duration.



#### 4. *The Pew and Gardner Procedure*

Pew and Gardner (8) also generated a confusion matrix using tachistoscopic exposure of upper case letters. Black letters, prepared with a Leroy lettering set, pen #4, templet #61 0250-425 C (approximately .425 inches high) on a white background were exposed one at a time with a Gerbrands two-channel tachistoscope. In contrast to the present experiment, the testing was conducted in a lighted room. The stimulus field was at full intensity (estimated to be approximately two ft.-L), and exposure duration (manipulated by superimposing a masking field of bits and pieces of letters) was varied from *S* to *S* and within *S*s according to an iterative technique which attempted to set the error rate for each *S* at about 50 per cent. The average error rate obtained by this technique for the 20 *S*s tested was actually 45.33 per cent and the average exposure duration was 29.72 msec with a range of 15-50 msec. Each *S* viewed each letter of the alphabet 10 times, thus each letter of the alphabet was viewed a total of 200 times.

#### 5. *The Hodge Procedure*

Hodge (5) generated a confusion matrix utilizing an altogether different procedure. Fifteen *S*s each viewed white stimulus cards each containing all 26 letters of the alphabet. The black upper case letters (approximately .24 inches high) were prepared with Leroy templet #3240-240 CL with Leroy pen numbers, 1-7. On each of 28 different cards the letters were arranged in a different random order, but were always presented in four rows of five letters each and one row of six letters. Four cards were prepared with each pen number.

The test cards and illuminating sources were mounted in an 18 by 18 by 36-inch flat-black enclosure. The luminance of the test card background was set at 25.8 ft.-L. The test cards were first shown to *S*s at a distance of 300 cm, and moved closer between trials in multiples of 10 cm; the number and size of steps varied among the *S*s. *S*'s task was to read off the letters on the test card as he recognized them. He was instructed to report a blank if he could not recognize a letter. The criterion employed was correct recognition of all 26 letters of the alphabet on two successive trials. This technique resulted in each letter being presented a total of 1218 times.

### C. RESULTS

The confusion frequencies obtained at exposure durations of 200 msec and 400 msec are presented in Tables 1 and 2 respectively. The analogous Pew

TABLE 1  
CONFUSION MATRIX: STIMULUS DURATION 200 MSEC

Ra	Stimulus letters													
	A	B	C	D	E	F	G	H	I	J	K	L	M	N
A		1	1		1	1					2	1	2	1
B				4	5	3	2	5	1	2	1	3	1	
C	1	2			3	2	3			1		6	1	
D		5	1		1		2		1	2		3		
E		1	1	1		1				1		5		
F	1		1	1	8				3	1	2	1		
G	1	6	10		1				1	3		4		
H		1			1	2			1		1		2	1
I	1	3		1	10	18		3		22	4	16	1	1
J	2	1			2	1		1	9			2	2	2
K	3	2	1		4	4		1	2					1
L		3	3	1	10	3	1	1	2	3	1			1
M	4					2		1			1	2		1
N	4	2				2	1	10	2	1	14	1	5	
O		3	14	17			13	2		3		2		
P	1	4				14			3		3	1	1	1
Q	1	1	2	3			3				1			
R	3	6			1	3		3	1	1	9			
S		4			1		2		1		1		1	
T			1	2	4	13	1		13	6	1	8		1
U	1	2		3	2		2	4		1		2		
V		2		1	1			2		1				
W			1											1
X	2				1				1					
Y					1	3		1	1	2	1	2	1	1
Z	3		1					1	2	1		1		
Total	28	49	37	34	57	72	30	35	44	51	42	59	17	12

and Gardner (8) data appear in Table 3; the Hodge (5) data in Table 4. The italicized cells in each matrix indicate confusions constituting five per cent or more of the number of presentations within that matrix.

In Figure 1, the per cent of total presentations of a stimulus letter that led to confusions have been plotted as a function of the stimulus-response pairs for each of the confusion matrices presented. In order to provide some indication of the extent to which similarities exist in the pattern of confusions across matrices all stimulus-response pairs that led to a five per cent level of confusions or better in Table 1 were rank ordered and plotted as shown. The corresponding levels of error for each stimulus-response pair in the other matrices were then plotted. Those letters in these latter three matrices which

TABLE 1 (*continued*)

Ra	Stimulus letters												Total
	O	P	Q	R	S	T	U	V	W	X	Y	Z	
A				2	1	1		1		5			20
B	1	1		4	13	1	2	1	1		1	2	54
C			2	1	3	1	2				2	2	31
D	1	1	1		3		4	1			2	2	30
E		2										4	16
F		3		2	1			2			2	2	30
G	4		7	1	2		2	1	1	1	1	1	47
H		3		4		1	2	1			1	1	22
I		11		1	1	35	1	1		4	6	3	143
J		2		1	2	5	6	1		2	4	1	46
K				11		3		1		20	2	4	59
L		1		3		4	3	1			2	4	47
M		2			1					4			17
N		1	1	7	3	1		1	1	6	1	1	65
O		1	30	1	6		11	4		1	3	1	112
P				5	2	2			1	1		1	40
Q	3			1								2	17
R		3	1		1	1		1		5		1	40
S		1		2		2				3	2	4	24
T		2	1	1	1		1			2	6	3	67
U	2			2						1	1		23
V		3		1			2		4	3	3	1	24
W		2	3			1		4		1			13
X	1	1		1		3					3	1	14
Y		1				3		15	4	3		5	44
Z								3		3			15
Total	13	41	46	51	40	63	36	39	12	64	42	46	1060

Note: Forty-one per cent of the presentations resulted in confusions. The italicized cells contain errors constituting five per cent or more of the total presentations of a letter.

<sup>a</sup> R = response.

exceeded the five per cent criterion and yet are not shown in Figure 1 are shown in Figure 2, rank ordered on the basis of the level of error obtained by Pew and Gardner (8).

Finally, to facilitate further the assessment of the similarity of the results obtained from study to study, the 15 stimulus-response pairs which resulted in the most confusions are shown in rank order for each matrix in Table 5. The letters used to prepare the table were identical to those utilized in each of the studies, except for the difference in size.

TABLE 2  
 CONFUSION MATRIX: STIMULUS DURATION 400 MSEC

R <sup>a</sup>	Stimulus letters													
	A	B	C	D	E	F	G	H	I	J	K	L	M	N
A		1				1		1				1	5	
B	2			3	5	1		3		1	1	1	1	
C		1			2		4					3		
D		1			1				1	1				
E		1				3				2		3	1	
F					5			1	1	1				
G	1	3	8		2					1				
H					2	2							1	2
I		1		1	2	18	1	2		24	1	17		2
J				1	1	3		1	5			2		
K		2												1
L		2	2	2	6	1			2	1				
M								6						
N		1				1		3	1		2	1	3	
O		2	4	6	1	1	10	1	2					
P	1	3			1	15					2			
Q	1		4				2							
R		2				3		2	1		3			
S		2			1		1							
T		1				5				7	2		3	
U				1					1	1				
V						1					1			
W								1		1	1		2	
X	1										1			
Y	2										2		1	
Z									1	1				
Total	8	23	18	14	29	55	18	21	15	41	16	28	17	5

## D. DISCUSSION

The two matrices stemming from the present study were generated under identical conditions except for a change in exposure duration. Increasing the exposure duration from 200 msec to 400 msec effectively reduced task difficulty as indicated by the decrease in total confusion errors from 41 per cent to 22 per cent. Of greater interest to the present study, however, is the extent to which the pattern of confusion errors changes as a function of exposure duration. It would appear from inspection of Figure 1 that a "best fit" curve for the 400 msec data would have the same general form as the 200 msec data (i.e., on the average the rank order position of letter pairs seems reasonably

TABLE 2 (continued)

R <sup>a</sup>	Stimulus letters												Total
	O	P	Q	R	S	T	U	V	W	X	Y	Z	
A				1						4	1	1	16
B		2		3	9		2						34
C	1				1	1	2					1	16
D	1	1			2		5					1	14
E							1	1				2	14
F		5			2	2					1	2	20
G				1	9		1				1	1	28
H				4		1	1					1	14
I		9		4		34	1			1	2		120
J				1		2	1	1	1		2	1	22
K				11						14	2	3	33
L						3	1	1				5	26
M									1			1	8
N		2		2		1			2	3	1		21
O	1	18		1	3		4			1	1		56
P			2	3				1			3		31
Q					2								9
R		2			1					2		1	17
S										1	1		6
T		1		1	1						6	1	28
U								1					4
V		1			1			2			2	1	9
W			1						2	1			9
X											1	1	4
Y		1		1	1			3		2			13
Z			1			1	1				1		6
Total	2	25	22	33	32	45	20	9	7	29	25	23	578

*Note:* Twenty-two per cent of the presentations resulted in confusions. The italicized cells contain errors constituting five per cent or more of the total presentations of a letter.

<sup>a</sup> R = response.

consistent). This process of averaging, however, is somewhat misleading because it eliminates local inconsistencies in the data. Assume, for example, that one wished to predict the relative confusability of two letter pairs under the 400 msec condition based on data collected at 200 msec. Looking at the 200 msec data one would predict that DO confusions occur more frequently than SB confusions. Examination of the 400 msec data, however, reveals the reverse to be true. Such inconsistencies, of course, tend to diminish as one deals with letter pairs that are more extremely separated on the curves. In short, while



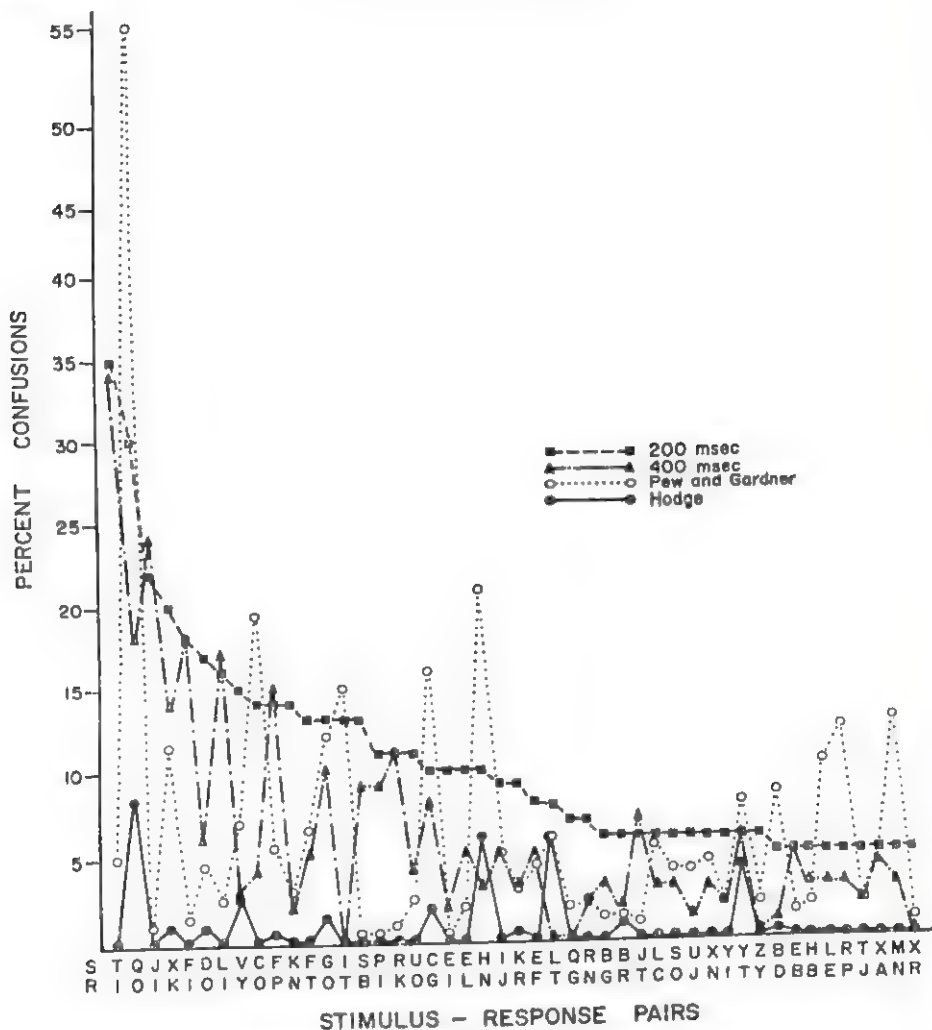


FIGURE 1

PER CENT OF TOTAL PRESENTATIONS OF EACH STIMULUS LETTER LEADING TO CONFUSIONS AS A FUNCTION OF STIMULUS-RESPONSE PAIRS AND STUDY

Only those stimulus-response pairs that led to at least a five per cent level of confusion under the 200 msec condition are included.

there is a general correspondence between the 200 msec and 400 msec matrices, predictions about specific confusions would seem to leave something to be desired.

When the results of the present study are compared with those of other investigators, the lack of similarity between matrices is even more apparent.

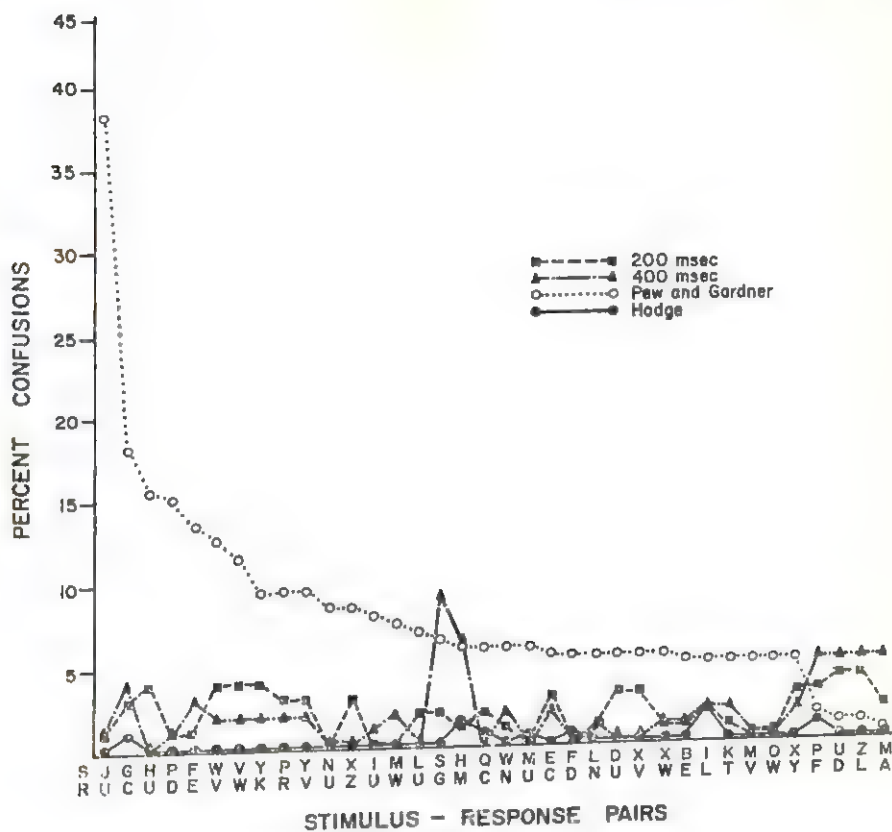


FIGURE 2

PER CENT OF TOTAL PRESENTATIONS OF EACH STIMULUS LETTER LEADING TO  
CONFUSIONS AS A FUNCTION OF STIMULUS-RESPONSE PAIRS AND STUDY

Only those stimulus-response pairs that led to at least a five per cent level of confusion and that are not shown in Fig. 1 are included.

In the case of the Pew and Gardner (8) study, task difficulty was essentially the same as the 200 msec condition of the present study (45 per cent and 41 per cent of the total presentations resulting in confusion, respectively). However, comparison of Figure 1 with Figure 2 and examination of Table 5 illustrate the almost total lack of correspondence between the patterns of error underlying these matrices. An equally discouraging result is obvious when one compares the present data with those of Hodge (5). The Hodge data fail to compare well with either of the matrices generated in the present study or with the data of Gardner and Pew.

Some of the differences between studies might be accounted for on the basis

TABLE 3  
CONFUSION MATRIX: AFTER PEW AND GARDNER (1965)<sup>a</sup>

R <sup>b</sup>	Stimulus letters												
	A	B	C	D	E	F	G	H	I	J	K	L	M
A		1		2	5	5	1	3	3	1	6	4	
B			1		3	2		4		2	4	4	1
C	1	6		5	11	6	36	5	5	2	6	11	
D	1	17	4		6	11		5	8	2	2	9	
E	2	10		1		27	1	5	2		2	21	
F	1			1	9			6	5		6	4	
G	1	3	32	1	5	3		3	1		2	2	1
H	2				3	3			7	2		7	6
I	1				1	3				2	3	5	
J	1	1	2	3	4	1		5	10			3	7
K	6	3			3	5		7	5			4	3
L				1	4	1		4	10		3		1
M		1				2		12	5	1	1	2	
N	2	1		1	6	5	1	42	8	2	6	11	26
O	2	4	39	9	1	3	24	1	3	2	1	4	
P		4	1	3	3	11	1	3	6	1	4	3	
Q			9	2		1	4	1		2			
R	2	3			8	9	2	3	6	1	6	4	1
S	4	6		1	6		2		2	3	5	2	
T					7	13		7	30	2	10	12	1
U		2	1	11	3	4	1	31	16	76	6	14	12
V	2	2		1	1	1			6	8	9	2	10
W				1	1	4		4	5	2	7	2	15
X	2				2	3		3		2	9	2	
Y					1	1		2	3		8		2
Z	2				9	3			2	2	4	5	1
Total	32	64	89	43	102	127	73	156	148	115	110	137	87

of differences in the styles of letters employed. For example, in both matrices generated from the present study, JI confusions ranked high, while they did not occur at all in the top 15 ranks of either the Pew and Gardner (8) or Hodge (5) study. It can be seen from Table 5 that the "deca dry" J has a very small tail relative to the Leroy J. This factor alone could presumably result in more confusions with I, especially since the Ss were not familiarized with the individual letters prior to their tachistoscopic exposure. Similar explanations can account for differences across studies between other pairs, such as TI where in the present study, the bar crossing the T was somewhat shorter than in earlier studies. In short then, some of the differences between the present study which used "deca dry" transfers and the previous studies both of

TABLE 3 (continued)

R <sup>b</sup>	Stimulus letters													Total
	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	
A	4		2		3		5	3	3	4	10	1	8	74
B			3	1	5	1		1		1	3	4	3	43
C	3	7	2	12	3	8	1	3	1	2	5	3	2	146
D	2		30	1	5	5	2	2	1		4	2	4	123
E			3		3	1	7	1		1		2	8	97
F	1		3		6		6	1			6	3	2	60
G	3	5	2	4	3	13	4	1	1		1	4	1	96
H	7		1		1		1	1	3	2	3	3		52
I	2		1		1		10	1				4	1	35
J	3	2					4	8	2	2	3	8	3	72
K	3		1		2		6		2	2	21	19	6	98
L	3						5			1	2	3	2	40
M	6				1		2		3	3	4	2	3	48
N					4		6	4	5	12	9	8	3	162
O			3	110	1	11		5	2	1	3	2	1	229
P	1				25		7		3	3	2	4		85
Q		8				5						2	1	35
R	1		19			1	4	1	1	2	2	5	7	88
S	2	1	4		4		3	1	2		2	3	5	58
T	2		4		2	1		1	2	3	8	16	3	124
U	17	1	1		1		5		2		6	2		212
V	4		1		1		3	2		25	11	19	3	111
W	10	2	1						23		11	6	6	100
X	4			1	1				1			1	5	36
Y	1				1	1	6	1	14	3	10		4	58
Z	4		2		2	2	7	1	4		17	8		75
Total	83	26	83	129	75	46	94	38	74	68	143	134	81	2357

Note: Forty-five per cent of the presentations resulted in confusions. The italicized cells contain errors constituting five per cent or more of the total presentations of a letter.

<sup>a</sup> Permission has been granted by Dr. Richard A. Pew and Gerald T. Gardner for the use of their unpublished data.

<sup>b</sup> R = response.

which used letters generated with a Leroy templet can be attributed to seemingly minute differences in style. Finally it should be remembered that the Hodge (5) Ss were allowed to report a blank if they did not recognize a letter, whereas Ss in the other two studies were required to identify every stimulus. This not only accounts for the somewhat lower confusion rate noted by Hodge but also may account for some of the differences in rank order of confusions.

TABLE 4  
 CONFUSION MATRIX: AFTER HODGE (1962)<sup>a</sup>

R <sup>b</sup>	Stimulus letters													
	A	B	C	D	E	F	G	H	I	J	K	L	M	N
A									1					
B					1			1						
C				3			14				1			
D		6	1				1							
E			1				1				4			
F		1			3		1			1				
G		1	23	1						1				
H		3											1	7
I														
J									3					
K	1	1			1			1					1	2
L	1		1		3		1	1	26					
M								16				1		
N		2						71					2	
O		2	3	12			16							
P				5		7		1						
Q			1	2			15							
R	1	11				2		4			8			2
S		33	1	3			41							2
T						3	1		2					
U		1					2	1					7	
V														
W		1						1					5	1
X							1				19			
Y					1	1								
Z	1	1			1						1			
Total	4	63	31	26	10	13	94	97	32	2	33	1	16	14

In summary then, there is little evidence for the common assumption that there exists a basic "pattern of confusions" between upper case letters of the alphabet. When tachistoscopic exposure and impoverished viewing conditions were used, a decrease in exposure duration resulted in the expected increase in task difficulty, but examination of the resulting matrices revealed only moderate similarity in the patterns of errors. On the other hand, when task difficulty was held constant while viewing conditions were modified (Pew and Gardner *versus* the present study), there was virtually no correspondence between the resulting pattern of confusion errors. Similarly, when the Hodge data were compared with those generated by either of the tachistoscopic techniques, correspondence was minimal.



TABLE 4 (*continued*)

R <sup>b</sup>	Stimulus letters												Total
	O	P	Q	R	S	T	U	V	W	X	Y	Z	
A				4			1						6
B	1			3	4			1					11
C	36		9										63
D	22		1	1	2						1		35
E										3			9
F		10				1		1					18
G	17		73		1								117
H				5					1				17
I	1												1
J							2			1			6
K										10			17
L					2								35
M				1					9				27
N				4		3			1				83
O			102	1	3		3						142
P											1		14
Q	7			1									26
R		1						1		1			31
S												1	81
T										1	56	1	64
U	1		1					8					21
V			1				1				5		7
W							1						9
X					1								21
Y						3		34	1	3			43
Z				1						5			10
Total	85	11	187	21	13	7	8	45	12	24	63	2	914

Note: Three per cent of the presentations resulted in confusions. The italicized cells contain errors constituting five per cent or more of the total presentations of a letter.

<sup>a</sup> Permission has been granted by the author and the publisher for the data which appeared on p. 42 of Hodge, D. C., Legibility of a uniform-strokewidth alphabet: I. Relative legibility of upper and lower case letters. *J. Eng. Psychol.*, 1963, 1, 34-46.

<sup>b</sup> R = response.

It appears, then, that confusion matrices are a function of the procedures and techniques by which they are generated. Considerable research will undoubtedly be required to isolate systematically all the pertinent variables involved. Such variables as exposure duration, report technique (i.e., forced *versus* free report), and letter style clearly warrant further study. In the interim it would appear that investigators of short-term memory and related phenomena wishing to make assumptions concerning the probability with which perceptual

TABLE 5  
RANK ORDER OF CONFUSIONS FOR THE FOUR MATRICES

RANK	Fisher 200		Fisher 400		Pew & Gardner		Hodge	
	S R	% con- fusion	S R	% con- fusion	S R	% con- fusion	S R	% con- fusion
1	T I	35	T I	34	Q O	55	Q O	84
2	Q O	30	J I	24	J U	38	Q G	6
3	J I	22	Q O	18	H N	21	H N	59
4	X K	20	F I	18	C O	19.5	Y T	4.5
5	F I	18	L I	17	G C	18	G S	34
6	D O	17	F P	15	C G	16	O C	28
7	L I	16	X K	14	H U	15.5	V Y	28
8	V Y	15	R K	11	I T	15	B S	28
9	C O	14	G O	10	P D	15	I L	2
10	F P	14	P I	9	F E	13.5	C G	2
11	K N	14	S B	9	M N	13	O D	1.8
12	F T	13	S G	9	R P	12.5	K X	1.6
13	G O	13	C G	8	W V	12.5	O G	1.5
14	I T	13	J T	7	G O	12	G O	1.5
15	S B	13	D O	6	V W	12	H M	1.5

confusions exist between letters of the alphabet must make their test procedures identical with those accompanying an existing confusion matrix or conduct preliminary investigations with procedures comparable to those to be used in subsequent efforts.

### E. SUMMARY

Perceptual confusion matrices were generated with the use of tachistoscopic exposure of upper case letters of the English alphabet at each of two exposure durations. The resulting matrices were compared with those generated by Hodge (5) and Pew and Gardner (8). Little correspondence was noted between the pattern of confusions obtained in each study. Thus, there was no evidence for the common assumption that a basic "pattern of confusions" exists between letters of the alphabet. Implications for studies of short-term memory were discussed.

### REFERENCES

1. CHASE, W. G., & POSNER, M. I. The effect of auditory and visual confusability on visual and memory search tasks. Paper presented at the Midwestern Psychological Association, Chicago, Illinois, May, 1965.

2. CIMBALO, R. S., & LAUGHERY, K. R. Short-term memory: Effects of auditory and visual similarity. *Psychon. Sci.*, 1967, 8, 57-58.
3. GIBSON, E., OSSER, H., SHIFF, W., & SMITH, J. An analysis of critical features of letters tested by a confusion matrix. From *A Basic Research Program on Reading*, Cooperative Research Project No. 639 (Final Report), Cornell University, Ithaca, New York, 1963.
4. GLUCKSBERG, S., FISHER, D. F., & MONTY, R. A. Brief visual memory as a function of visual and acoustic confusability. Proceedings 75th Annual Convention, American Psychological Association, Washington, D.C., 1967, 55-56.
5. HODGE, D. C. Legibility of a uniform-strokewidth alphabet: I. Relative legibility of upper and lower case letters. *J. Eng. Psychol.*, 1962, 1, 34-46.
6. KAPLAN, G. A., YONAS, A., & SHURCLIFF, A. Visual and acoustic confusability in a visual search task. *Percept. & Psychophys.*, 1966, 1, 172-174.
7. KUNNAPAS, T. Visual perception of capital letters: Multidimensional ratio scaling and multidimensional similarity. *Scand. J. of Psychol.*, 1966, 7, 189-196.
8. PEW, R. W., & GARDNER, G. T. Unpublished data, University of Michigan, Ann Arbor, Michigan, 1965.

*Behavioral Research Laboratory*  
*U.S. Army Human Engineering Laboratories*  
*Aberdeen Proving Ground, Maryland 21005*



## FURTHER STUDY OF PERFORMANCE ERRORS ON RAVENS PROGRESSIVE MATRICES (1938)\*<sup>1</sup>

*Research Division, Central Islip State Hospital*

---

CHARLES SHEPPARD, DIANE FIORENTINO, LOIS COLLINS,  
AND SIDNEY MERLIS

---

### A. INTRODUCTION

Significant differences in reasoning errors on the Ravens Progressive Matrices (RPM) were reported between sociopathic and schizotypic drug users. Performance of schizotypic patients was more erratic; they committed more avoidable errors than sociopathic patients (12).

This study seeks to extend previous work by examining (RPM) performance of three additional groups of patients.

### B. METHOD

Within five days of their admission to the narcotic unit at Central Islip State Hospital, a short social history form and a battery of psychological tests was offered each patient. In an effort to ensure validity of patient response, no one was coerced into testing.

The social history form was designed to obtain biographical, educational, and occupational experience, as well as history of drug use. The psychological test battery included the Ravens Progressive Matrices (7), Institute of Personality and Ability Testing, Eight Parallel-Form Anxiety Battery Form-A, IPAT-8-A (8) and Minnesota Multiphasic Personality Inventory, MMPI (2).

Data were gathered from 396 patients committed to the unit by a probate court procedure for a 90-day treatment period.<sup>2</sup> The characteristics of the first hundred patients have been reported elsewhere and can serve as comparative information, since the characteristics of the population have remained relatively stable throughout the testing period (9, 13).

---

\* Received in the Editorial Office, Provincetown, Massachusetts, on November 4, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This study was supported in part by Public Health Service Grant No. MH-05096-06 from the National Institute of Mental Health.

<sup>2</sup> The authors thank Francis J. O'Neill and Leslie Eber, M.D. for their cooperative efforts in making the patients available for testing. They thank Geraldine Burkie, M.A., Miss Lucy Lemmer, Miss Sue Rosenberg and Rita Kaplan, M.A., for their assistance in testing the patients.



All MMPI profile sheets were separated from the patients' record folders, given an identification number, then sorted into groups by profile pattern. Objective configural rules were observed in classifying MMPI records (1, 5, 6). Three subtypes of patients were chosen for further analysis, those defined by a (42') secondary sociopathic pattern ( $N = 20$ ), a (428') paranoid pattern ( $N = 19$ ), and a (987') schizophrenic-mixed pattern ( $N = 15$ ). Description of the patients sampled indicates reliability of the profile assignment. Analysis of variance established significant  $F$  ratios, at the .01 level, on seven MMPI scales (see Table 1).

Subsequent  $t$  tests demonstrated that the secondary sociopathic group scored significantly higher than the schizophrenic-mixed group on the K ( $t = 3.04$ ), ( $p < .01$ ), Pd ( $t = 2.04$ ,  $p = .05$ ), and D ( $t = 2.78$ ,  $p < .01$ ); whereas the schizophrenic-mixed group scored higher on the Pa ( $t = 5.43$ ,  $p < .001$ ), Pt ( $t = 4.27$ ,  $p < .001$ ) and Ma ( $t = 7.38$ ,  $p < .001$ ) scales. Significant mean differences  $t$  were calculated between the paranoid and schizophrenic-mixed groups.

The paranoid group scored significantly higher on the K ( $t = 3.04$ ,  $p < .01$ ), D ( $t = 2.78$ ,  $p < .01$ ), Pd ( $t = 3.63$ ,  $p < .01$ ); whereas the schizophrenic-mixed group scored higher on the Pt ( $t = 2.94$ ,  $p < .01$ ) and Ma ( $t = 5.40$ ,  $p < .001$ ) scales. Significant mean differences were calculated between the secondary sociopathic and paranoid groups.

The paranoid group scored significantly higher on the Pa ( $t = 5.37$ ,  $p < .001$ ), Sc ( $t = 4.27$ ,  $p < .001$ ), but significantly lower on the Ma ( $t = 2.17$ ,  $p < .05$ ) scale than the secondary sociopaths. The data demonstrated statistical differences between the three MMPI profile types. No significant differences were noted among the groups in RPM centile scores, age, duration of heroin use, years of formal education, or IPAT-8-A centile scores (see Table 1). With the definition of the sample established, patients' RPM answer sheets were separated from their records, given an identification number, then scored for avoidable errors.

To establish whether an error was avoidable, it was necessary to determine if the patient possessed the ability to solve the item. In accordance with the RPM design and empirically established item difficulty (10), patients missing an item then solving other items of greater difficulty have the reasoning ability to solve the easier item. The lowest difficulty ( $p$  level) the patient consistently attained was taken as the best estimate of his reasoning capacity. This cutoff score was used as a point of departure in scoring avoidable errors. Items missed with higher  $p$  level (easier items) were considered avoidable errors, since they required less conceptualization to solve them. Thus each patient was used as

TABLE 1  
MEAN AGE, EDUCATION, YEARS USING HEROIN, RPM AND IPAT CENTILES AND MMPI T SCORES, STANDARD DEVIATION  
AND F VALUES FOR THREE MMPI PROFILE TYPES

Variable	42' pattern (N=20) <sup>a</sup> $\bar{X}$	$\sigma$	428' pattern (N=19) <sup>b</sup> $\bar{X}$	$\sigma$	987' pattern (N=15) $\bar{X}$	$\sigma$	F	p
Age	30 Yrs.	6.43	29 Yrs.	5.52	30 Yrs.	6.02	.11	NS
Heroin use	9 Yrs.	4.62	8 Yrs.	3.80	11 Yrs.	4.67	.95	NS
Education	10 Yrs.	2.60	9 Yrs.	2.03	9 Yrs.	2.09	.08	NS
Ravens centile	48	29.60	51	25.83	41	30.00	.54	NS
IPAT centile	38	28.45	29	26.05	32	29.04	.50	NS
L	48	6.47	51	7.90	48	9.13	1.06	NS
F	53	11.81	54	23.35	46	27.43	.65	NS
K	53	7.95	51	8.32	44	8.89	5.17	.01
Hypochondriasis	55	5.48	55	6.83	52	12.68	.72	NS
Depression	71	6.73	73	9.69	62	10.73	6.41	.01
Hysteria	58	7.06	57	5.31	52	7.18	3.09	NS
Psychopathic deviate	76	7.32	81	8.05	70	10.01	6.63	.01
Masculinity-femininity	58	7.56	59	9.99	63	8.24	1.85	NS
Paranoid	51	8.16	67	9.52	64	6.30	20.62	.01
Psychasthenia	59	7.21	61	8.08	72	12.56	10.25	.01
Schizophrenia	55	4.88	71	7.80	73	20.17	13.18	.01
Hypomania	58	8.19	64	6.30	79	10.20	29.06	.01
Social introversion	52	6.67	57	6.72	56	8.22	2.36	NS

<sup>a</sup> The authors wish to thank Fredrick C. Thorne, M.D., Ph.D., Editor, of *The Journal of Clinical Psychology* for permission to reproduce portions of test data soon to be published in Sheppard, Fiorentino, Collins, and Merlis, "Comparison of emotion profiles as defined by two additional MMPI profile types in male narcotic addicts" (11).

<sup>b</sup> The authors wish to thank Carol Ammons, Ph.D., Editor, *Psychological Reports*, for permission to reproduce portions of the test data published in Sheppard, Fiorentino, and Merlis, "The affective differential: A comparison of emotion profiles gained from clinical judgment and patient self-report," 1968, 22, 809-814 (9).

his own control. A failure to solve an item whose difficulty level was within the testee's range of ability, as determined by his performance, was scored as an avoidable error.

Reliability and construct validity of the scoring technique were defined by product moment correlation coefficients measuring the interrelationship between RPM centile score, lowest item difficulty level, and number of avoidable errors for each group separately. Tests of significance  $t$  were computed to test the mean differences in the number of avoidable errors committed by each profile group.

### C. RESULTS

Before evaluating the frequency of avoidable errors committed by the different profile pattern groups, it was necessary to demonstrate the reliability of the technique for scoring avoidable errors. Estimate of reliability may derive from internal consistency with regard to measures of association between RPM centile and item difficulty cutoff scores and between RPM centile and number of avoidable errors across the three profile groups.

Product moment correlation coefficients measuring the relationship between patients' RPM centile score and item difficulty level equalled  $-.88$  for the secondary sociopaths,  $-.85$  for the paranoid group, and  $-.88$  for the schizophrenic-mixed group. These high negative correlations were significant from zero at less than the  $.001$  level. Tests of significance  $t$  between pairs of correlations demonstrated no difference. The consistency of the relationship across profile pattern indicated patients achieving higher RPM centile scores solve the more difficult test items regardless of type or degree of illness.

Consider also the relationship between RPM centile scores and number of avoidable errors. Product moment coefficients equalled  $-.15$  for the secondary sociopaths and  $-.06$  for the paranoid and schizophrenic-mixed groups. The low, negative correlations were not significant from zero, suggesting no relationship between RPM centile score and number of avoidable errors. Tests of significance  $t$  between pairs of correlations demonstrated no difference among the groups. Combined data show avoidable errors were scored when a patient failed an item that was within his capacity to solve. The consistency of the relationship across groups argued for the reliability of the empirical procedure for scoring avoidable errors. With the reliability of the procedure of scoring avoidable errors defined, a test of the hypothesis of no difference between the three profile types was undertaken.

The mean number of avoidable errors, the standard errors of the means, and  $t$  values are seen in Table 2. Both the secondary sociopathic and schizo-

phrenic-mixed groups committed significantly fewer avoidable errors than the paranoid group. The hypothesis that type and degree of emotional disturbance result in differences in the frequency of avoidable errors was supported.

TABLE 2  
MEAN NUMBER OF AVOIDABLE ERRORS, STANDARD ERRORS OF THE MEANS,  
 $t$  VALUES AND  $N$  FOR THE THREE MMPI PROFILE TYPES

Statistic	42' pattern	428' pattern	987' pattern
$\bar{X}$	3.1	5.5	3.4
$SE_m$	.59	.73	.63
$N$	20	19	15

Note: 42'/428',  $t = 2.62$ ,  $p = .05$ ; 42'/987',  $t = .41$ , N.S.; 428'/987',  $t = 2.18$ ,  $p = .05$ .

#### D. DISCUSSION

In an earlier study patients defined as sociopathic and schizotypic, on the basis of MMPI profile patterns, were shown to differ significantly in the numbers of reasoning errors made on the RPM (12). These data should be considered exploratory in view of the past history of negative findings (3, 4). The present data tend to confirm and extend previous work.

Patients classified with 42' MMPI patterns may be considered as "secondary sociopathic" types. They differ from patients classified by 49' MMPI patterns, or "primary sociopaths," in their depression, hysteria, and lower hypomania scores. These groups are highly similar in the number of avoidable errors committed. Mean number of avoidable errors equalled 3.4 for primary sociopaths and 3.1 for the secondary sociopathic group.

Patients classified with 428' patterns may be considered paranoid personality types. The 987' pattern combines elements of manic, schizophrenic, anxious behavior and may be considered as schizophrenic-mixed type. It was expected that these groups might be similar in their performance to the earlier studied psychotic group. This expectation was not confirmed. The schizophrenic-mixed patients were similar to the sociopathic groups ( $\bar{X} = 3.4$ ), while the schizotypic and paranoid patients were similar ( $\bar{X} = 5.1$  and 5.5) in their performance. Further work is in progress to examine these differences.

#### E. SUMMARY

In an attempt to identify the existence of reasoning errors in psychotic patients, the performance of male narcotic users defined as secondary sociopaths ( $N = 20$ ), paranoid ( $N = 19$ ), and schizophrenic-mixed type ( $N = 15$ )

by MMPI profile type was analyzed for avoidable errors on the Ravens Progressive Matrices. An avoidable error was defined as a failure to solve an item whose difficulty level was within the testee's range of ability to solve as measured by his performance.

Construct validity was demonstrated with regard to intercorrelations of RPM centile score, item difficulty levels, and number of avoidable errors. Tests of significance  $t$  showed paranoid patients committed more avoidable errors than the secondary sociopathic and mixed-schizophrenic patients, but no difference was noted between the mixed-schizophrenic and secondary sociopathic patients.

### REFERENCES

1. GILBERSTADT, H., & DUKER, J. A Handbook for Clinical and Actuarial MMPI Interpretation. Philadelphia, Pa.: Saunders, 1956.
2. HATHAWAY, S. R., & MCKINLEY, J. Minnesota Multiphasic Personality Inventory Manual. New York: Psychological Corp., 1951.
3. JURJEVICH, R. M. Avoidable errors on Raven Progressive Matrices and psychopathological indices. *Psychol. Rep.*, 1967, **21**, 364.
4. KASPER, S. Progressive Matrices (1938) and emotional disturbance. *J. Consult. Psychol.*, 1958, **23**, 1.
5. MARKS, P. A., & SEEMAN, W. The Actuarial Description of Abnormal Personality. Baltimore, Md.: Williams & Wilkins, 1963.
6. MEEHL, P. E., & DAHLSTROM, W. G. Objective configural rules for discriminating psychotic from neurotic MMPI profiles. *J. Consult. Psychol.*, 1960, **24**, 375-387.
7. RAVEN, J. C. Guide to the Standard Progressive Matrices, A, B, C, D, & E. London, England: H. K. Lewis, 1960 (distr. in U.S. by Psychological Corp.).
8. SCHEIER, I. H., & CATTELL, R. B. Handbook and Kit for the IPAT-8-Parallel Form Anxiety Battery. Champaign, Ill.: Inst. Personal. & Ability Testing, 1960.
9. SHEPPARD, C., FIORENTINO, D., & MERLIS, S. Affective differential: Comparison of emotion profiles gained from clinical judgments and patient self-report. *Psychol. Rep.*, 1968, **22**, 809-814.
10. SHEPPARD, C., FIORENTINO, D., COLLINS, L., & MERLIS, S. Ravens Progressive Matrices (1938): Normative data on male narcotic addicts. *Psychol. Rep.*, 1968, **23**, 343-348.
11. ———. Comparison of emotion profiles as defined by two additional MMPI profile types in male narcotic addicts. *J. Clin. Psychol.*, in press.
12. ———. Performance errors on Ravens Progressive Matrices (1938) between sociopathic and schizotypic personality types. Submitted to *Psychol. Rep.*, 1968, **23**, 1043-1046.
13. SHEPPARD, C., LEMMER, L., O'NEILL, F. J., & MERLIS, S. A study of one hundred admissions to the Narcotic Unit at Central Islip State Hospital: Some preliminary inferences drawn from self-report questionnaires. Presented at the Department of Mental Hygiene Downstate Interhospital Conference, Upstate Medical Center, Syracuse, New York, April 25, 1966.

Research Division

Central Islip State Hospital

Central Islip, New York 11722



## COMPARISON BETWEEN AMERICAN PERMANENT RESIDENTS OF ISRAEL: PART I, AMERICAN BACKGROUND\*<sup>1</sup>

*B'nai B'rith Hillel Foundation at Purdue University*

---

GERALD ENGEL

---

### A. INTRODUCTION

Comfortable American Jews do not settle in numbers in Israel. They remain within the structure of the American community (9, 16, 24). American Jews living in Israel observed their own homeostatic adjustment to life in America was gradually overcome by a need for personal fulfillment in Israel (13, 15, 26). American students at Hebrew University are similarly drawn to Israel (8).

Today the Jew in the United States and Canada is seldom depicted as a marginal man struggling to adjust to the majority's culture (25). This minority has benefitted from cross-fertilization (6).

Nevertheless, members of this accommodating cultural minority still experience disturbing dissonance (12), as well as creative dissonance (10, 11). Acculturation has simultaneously reduced the fear of anti-Semitism while increasing intermarriage (17, 18, 19, 20, 21), assimilation of materialistic ideals, and racial prejudice (1, 4).

Whether those who leave America for Israel are more concerned about some areas of intergroup and interpersonal relations is explored in this study of native American Jews now living in Israel who became permanent residents during the five years, 1962 through 1966. During consideration of their American background, a basis is provided for observing to what extent American *aliya* is motivated by a reaction to the majority culture (14), as well as by

---

\* Received in the Editorial Office, Provincetown, Massachusetts, on November 4, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> The author acknowledges the aid of Dr. Harold Kahn, National Institute of Health, in preparing the questionnaire; of Dr. Ben Winer, Purdue University, for statistical analysis; Glenda McCracken and Arnold Woodruff, Purdue Computer Center, for programming. Marilyn Kopelowitz Engel was study coordinator.

Thanks are expressed to the Governing Board of B'nai B'rith Hillel Foundations for granting the author a decennial leave to gather the data from November 1966 through August 1967, and to Dr. Frederick L. Hovde, President of Purdue University, for allocating President's funds for computer work. B'nai B'rith generously provided funds for immediate publication.

a personal need for Jewish identification and "moral affirmation" (2) in Israel.

### B. SUBJECTS

From an official list of permanent residents of Israel the names of 780 Americans born in the United States or Canada were selected, comprising all native Americans who became permanent residents of Israel during the years 1962 through 1966, and who were 20 years of age or older by the end of 1966. Seventy-seven subjects were eliminated, including six who said they left America as infants, six who were deceased, and 65 who were not in Israel from

TABLE 1  
BIRTHPLACE, FAMILY BACKGROUND, AND AGE OF ARRIVAL IN ISRAEL<sup>a</sup>

Question	Percentages	
	Nonreligious <i>N</i> = 188	Religiously oriented <i>N</i> = 255
Place of birth:	( <i>N</i> = 188)	( <i>N</i> = 255)
U.S.A.	89.4	91.4
Canada	10.6	8.6
Father born in:	( <i>N</i> = 188)	( <i>N</i> = 253)
America	30.3	25.3
Israel	1.1	1.6
Other	68.6	73.1
Mother born in:	( <i>N</i> = 187)**	( <i>N</i> = 253)
America	43.3	31.6
Israel	1.1	3.6
Other	55.6	64.8
Home was religiously:	( <i>N</i> = 186)	( <i>N</i> = 248)***
Orthodox	13.4	62.1
Conservative	26.9	24.2
Reform	11.3	3.6
Culturally Jewish	32.8	8.5
Other	15.6	1.6
Home was Zionistically:	( <i>N</i> = 183)	( <i>N</i> = 249)***
anti-Zionist	3.3	2.4
neutral	58.5	29.3
Zionist	38.3	68.3
First trip to Israel at age:	( <i>N</i> = 188)	( <i>N</i> = 255)*
under 24	50.0	59.6
25-34	26.1	26.3
35 & over	23.9	14.1

<sup>a</sup> All tables derived from BMD02s—Contingency Table Analysis—Version of March 11, 1964, Health Sciences Computing Facility, University of California at Los Angeles.

\*  $\chi^2 = .025$ .

\*\*  $\chi^2 = .02$ .

\*\*\*  $\chi^2 = .001$ .

March through August 1967 when this questionnaire was administered. Of 703 potential subjects 443 cooperated, representing a 63 per cent response.

### C. PROCEDURE

A brief statement at the beginning of the four-page self-administrating questionnaire advised recipients that native Americans in Israel were receiving the questionnaire as part of an independent research project to increase understanding of American *aliya* (immigration). Subjects were assured responses would be processed anonymously.

Subjects were asked to check biographical background, relation to parental

TABLE 2  
AMERICAN SECULAR AND JEWISH SCHOOLING

Question	Percentages	
	Nonreligious <i>N</i> = 188	Religiously oriented <i>N</i> = 255
School years completed:	( <i>N</i> = 188)	( <i>N</i> = 254)
0-12 (high school)	21.3	22.0
13-16 (college)	47.3	44.1
17-20 (post graduate)	31.4	33.9
Technical school:	( <i>N</i> = 176)	( <i>N</i> = 226)
no	88.1	90.7
yes	11.9	9.3
Sunday school:	( <i>N</i> = 147)**	( <i>N</i> = 231)
no	58.5	74.5
yes	41.5	25.5
After public school learning:	( <i>N</i> = 147)	( <i>N</i> = 232)
no	50.3	54.3
yes	49.7	45.7
Day school (parochial school):	( <i>N</i> = 148)	( <i>N</i> = 232)***
no	89.2	46.6
yes	10.8	53.4
Jewish graduate studies:	( <i>N</i> = 183)	( <i>N</i> = 239)*
no	95.1	88.3
yes	4.9	11.7
Hebrew teachers seminary:	( <i>N</i> = 177)	( <i>N</i> = 230)***
no	94.4	72.6
yes	5.6	27.4
Rabbinical seminary:	( <i>N</i> = 182)	( <i>N</i> = 241)***
no	97.8	85.1
yes	2.2	14.9

Note: 9.5 per cent had no Jewish schooling at all (derived by sorting).

\*  $\chi^2 = .02$ .

\*\*  $\chi^2 = .005$ .

\*\*\*  $\chi^2 = .001$ .

home, schooling, participation in Jewish organizations. A checklist included what they did not like about America, as well as Israel's attraction.

Personal feelings upon arrival, as well as comparisons between Israel and America, were solicited. Checklists included factors that would influence Americans to come, to stay, and to leave Israel. Opinions about the future of the Jew in America and the respondent's personal sense of satisfaction with living accommodations and life in Israel were also solicited. Space was allotted at the close of page four for remarks.

TABLE 3  
JEWISH ORGANIZATION PARTICIPATION AND RELIGIOUS IDENTIFICATION

Question	Percentages	
	Nonreligious <i>N</i> = 188	Religiously oriented <i>N</i> = 255
Participated in Jewish organizations:		
Zionist youth	( <i>N</i> = 164)	( <i>N</i> = 214)***
no	57.9	31.3
yes	42.1	68.7
Zionist adult	( <i>N</i> = 151)	( <i>N</i> = 165)***
no	78.1	58.8
yes	21.9	41.2
Synagogue	( <i>N</i> = 148)	( <i>N</i> = 172)***
no	70.3	44.2
yes	29.7	55.8
Hillel	( <i>N</i> = 149)	( <i>N</i> = 177)**
no	81.9	70.6
yes	18.1	29.4
Others	( <i>N</i> = 111)	( <i>N</i> = 105)*
no	82.9	72.4
yes	17.1	27.6
Consider self in America:	( <i>N</i> = 185)	( <i>N</i> = 254)***
Orthodox	2.2	57.9
Conservative	22.7	25.6
Reform	13.5	3.1
Culturally Jewish	38.9	11.4
Other	22.7	2.0

\*  $\chi^2 = .10$  trend.

\*\*  $\chi^2 = .02$ .

\*\*\*  $\chi^2 = .001$ .

Subjects were mailed a printed letter soliciting their cooperation, the self-administrating questionnaire, and a stamped, self-addressed return envelope, on March 8. The second form letter, questionnaire, and return envelope were sent March 22. A third mailing was made on May 3. Following the Six Day War in June, a handwritten postscript was added to the original

covering letter requesting participation in order to help encourage immigration to Israel (*aliya*). The response to this final appeal equalled the response to the third request. The 364 respondents to these four mailings are considered regular subjects.

Forced subjects included two groups. The first were Americans who had not responded to the initial request though their spouses had. These recalcitrants did respond to personal notes requesting equal cooperation. The larger group consisted of subjects from Jerusalem, Rehovot, Holon, Bat Yam, and Ramat Gan who did not answer the first two requests. They were advised by a handwritten postscript on the third letter, "Please save us the time and expense of a personal visit."

Several subjects responded by mail and were considered forced. The majority were interviewed by three Americans (Helen Bragman, and Marilyn and Gerald Engel) and two Israeli who had lived in America (Hanya Aharoni and Shula Eshel). Most interviewees were apologetic and cooperative. Only 14 persons refused to cooperate for personal reasons.

The responses of all 79 forced subjects were compared to those of the 364 regular respondents; few differences were observed. The two groups were then combined.

These 443 subjects were divided into two groups based upon their acceptance or rejection of Judaism as a religion. In order to ascertain the subjects' current religious views, six choices were presented in Hebrew. Two hundred fifty-five responded to the inquiry, "In Israel I consider myself," by identifying with Jewish religious tradition. Of these, 87 asserted they were traditionalists (*masorti*), 30 were keepers of commandments (*shomer mitzvot*), while 138 were observant (*dati*). The 188 who answered negatively were, for purposes of comparison, designated as nonreligious. These included 27 against observance (*anti-dati*), 80 not observant (*lo dati*), 55 secularists (*chiloni*), 24 who wrote "don't know," as well as two who declared they were Hebrew-Christians.

#### D. SOME HYPOTHESES

Hypothesis 1. The religiously oriented identified with American Jewish life more than the nonreligious did.

Hypothesis 2. The religiously oriented were interested in Israel more than the nonreligious were.

Hypothesis 3. The religiously oriented were disturbed by aspects of American life more than the nonreligious were.



home, schooling, participation in Jewish organizations. A checklist included what they did not like about America, as well as Israel's attraction.

Personal feelings upon arrival, as well as comparisons between Israel and America, were solicited. Checklists included factors that would influence Americans to come, to stay, and to leave Israel. Opinions about the future of the Jew in America and the respondent's personal sense of satisfaction with living accommodations and life in Israel were also solicited. Space was allotted at the close of page four for remarks.

TABLE 3  
JEWISH ORGANIZATION PARTICIPATION AND RELIGIOUS IDENTIFICATION

Question	Percentages	
	Nonreligious <i>N</i> = 188	Religiously oriented <i>N</i> = 255
Participated in Jewish organizations:		
Zionist youth	( <i>N</i> = 164)	( <i>N</i> = 214)***
no	57.9	31.3
yes	42.1	68.7
Zionist adult	( <i>N</i> = 151)	( <i>N</i> = 165)***
no	78.1	58.8
yes	21.9	41.2
Synagogue	( <i>N</i> = 148)	( <i>N</i> = 172)***
no	70.3	44.2
yes	29.7	55.8
Hillel	( <i>N</i> = 149)	( <i>N</i> = 177)**
no	81.9	70.6
yes	18.1	29.4
Others	( <i>N</i> = 111)	( <i>N</i> = 105)*
no	82.9	72.4
yes	17.1	27.6
Consider self in America:	( <i>N</i> = 185)	( <i>N</i> = 254)***
Orthodox	2.2	57.9
Conservative	22.7	25.6
Reform	13.5	3.1
Culturally Jewish	38.9	11.4
Other	22.7	2.0

\*  $\chi^2 = .10$  trend.

\*\*  $\chi^2 = .02$ .

\*\*\*  $\chi^2 = .001$ .

Subjects were mailed a printed letter soliciting their cooperation, the self-administrating questionnaire, and a stamped, self-addressed return envelope, on March 8. The second form letter, questionnaire, and return envelope were sent March 22. A third mailing was made on May 3. Following the Six Day War in June, a handwritten postscript was added to the original

covering letter requesting participation in order to help encourage immigration to Israel (*aliya*). The response to this final appeal equalled the response to the third request. The 364 respondents to these four mailings are considered regular subjects.

Forced subjects included two groups. The first were Americans who had not responded to the initial request though their spouses had. These recalcitrants did respond to personal notes requesting equal cooperation. The larger group consisted of subjects from Jerusalem, Rehovot, Holon, Bat Yam, and Ramat Gan who did not answer the first two requests. They were advised by a handwritten postscript on the third letter, "Please save us the time and expense of a personal visit."

Several subjects responded by mail and were considered forced. The majority were interviewed by three Americans (Helen Bragman, and Marilyn and Gerald Engel) and two Israeli who had lived in America (Henya Aharoni and Shula Eshel). Most interviewees were apologetic and cooperative. Only 14 persons refused to cooperate for personal reasons.

The responses of all 79 forced subjects were compared to those of the 364 regular respondents; few differences were observed. The two groups were then combined.

These 443 subjects were divided into two groups based upon their acceptance or rejection of Judaism as a religion. In order to ascertain the subjects' current religious views, six choices were presented in Hebrew. Two hundred fifty-five responded to the inquiry, "In Israel I consider myself," by identifying with Jewish religious tradition. Of these, 87 asserted they were traditionalists (*masorti*), 30 were keepers of commandments (*shomer mitzvot*), while 138 were observant (*dati*). The 188 who answered negatively were, for purposes of comparison, designated as nonreligious. These included 27 against observance (*anti-dati*), 80 not observant (*lo dati*), 55 secularists (*chiloni*), 24 who wrote "don't know," as well as two who declared they were Hebrew-Christians.

#### D. SOME HYPOTHESES

Hypothesis 1. The religiously oriented identified with American Jewish life more than the nonreligious did.

Hypothesis 2. The religiously oriented were interested in Israel more than the nonreligious were.

Hypothesis 3. The religiously oriented were disturbed by aspects of American life more than the nonreligious were.

Hypothesis 4. The religiously oriented came to live in Israel in fewer trips and less years from the initial trip than the nonreligious did.

### E. RESULTS

Hypothesis 1 is confirmed by comparing religious and Zionist home background (Table 1), Jewish schooling (Table 2), and participation in American Jewish community life (Table 3).

Hypothesis 2 is confirmed by comparing Zionist home background, noting earlier attraction to Israel (Table 1), and identification with Zionist groups (Table 3). Ability to speak Hebrew before arrival indicates interest in the land as does direct immigration to Israel (Table 4).

TABLE 4  
FACTS CONCERNING EMIGRATION

Question	Percentages	
	Nonreligious <i>N</i> = 188	Religiously oriented <i>N</i> = 255
Lived in other lands	( <i>N</i> = 179)*	( <i>N</i> = 250)
no	84.9	93.6
yes	15.6	6.4
Came to live in Israel on:	( <i>N</i> = 187)	( <i>N</i> = 254)
first trip	39.6	43.3
second trip	41.7	42.5
third trip or more	18.7	14.2
Years after first trip came to live in Israel:	( <i>N</i> = 186)	( <i>N</i> = 251)
on first trip	40.3	44.2
1-4	35.5	30.7
5 or more	24.2	25.1
On arrival in Israel, ability in conversational Hebrew	( <i>N</i> = 187)	( <i>N</i> = 251)**
none	54.5	20.7
some	26.2	27.9
much	19.3	51.4

\*  $\chi^2 = .005$ .

\*\*  $\chi^2 = .001$ .

Hypothesis 3 is confirmed only regarding assimilation, which relates to group survival. The nonreligious were more disturbed about the freedom of the individual than were the religious. Conformity and witchhunting disturb those who more readily find their place in the community at large (Table 5).

Hypothesis 4 is rejected. The religiously oriented do not come to live in

Israel in fewer trips and less years from their initial trip than the nonreligious (Table 4).

TABLE 5  
DISTURBING FACTORS ABOUT AMERICAN LIFE

Question	Percentages	
	Nonreligious <i>N</i> = 188	Religiously oriented <i>N</i> = 255
Materialism	( <i>N</i> = 165)	( <i>N</i> = 204)
none	26.1	29.4
some	37.6	33.3
much	36.4	37.3
Conformity	( <i>N</i> = 163)*	( <i>N</i> = 198)
none	23.3	31.3
some	30.1	35.4
much	46.6	33.3
Political witchhunting	( <i>N</i> = 158)**	( <i>N</i> = 189)
none	41.8	58.7
some	29.7	26.5
much	28.5	14.8
Church-state relationships	( <i>N</i> = 156)	( <i>N</i> = 186)
none	67.9	71.5
some	25.0	22.0
much	7.1	6.5
Assimilation	( <i>N</i> = 160)	( <i>N</i> = 200)***
none	43.8	27.0
some	34.4	30.5
much	21.9	42.5
Anti-Semitism	( <i>N</i> = 162)	( <i>N</i> = 203)
none	26.5	27.1
some	57.4	49.8
much	16.0	23.2
Educational system	( <i>N</i> = 156)	( <i>N</i> = 190)
none	67.3	67.9
some	25.6	24.7
much	7.1	7.4
Dating and marriage	( <i>N</i> = 157)	( <i>N</i> = 194)
none	57.3	57.2
some	29.3	25.3
much	13.4	17.5
Dependence on family	( <i>N</i> = 158)	( <i>N</i> = 193)
none	70.3	73.6
some	22.8	21.8
much	7.0	4.7

\*  $\chi^2 = .05$ .

\*\*  $\chi^2 = .005$ .

\*\*\*  $\chi^2 = .001$ .

While the two groups differ in the strength of their Jewish identification, the nonreligious who do not identify as strongly do identify to a greater extent than their counterparts in America.

Of all subjects only 9.5 per cent had no Jewish schooling (Table 2). Two thirds could converse in Hebrew upon arrival (Table 4). An equal proportion affiliated with Zionist groups (Table 3). The preponderance of second generation Americans (3) among these subjects indicates most Americans in Israel are products of immigrant homes where Jewish identification is taken for granted (Table 1). Nine-tenths of those motivated to settle in Israel went directly to this land (Table 1). Yet, the desire for Jewish identification and "moral affirmation" in a Jewish state is tempered by a desire to get away from a "sick society" as viewed not only by prevailing fear of anti-Semitism but also by considerable concern regarding materialism and conformity (Table 5) though the high level of secular education might indicate (Table 2) a high degree of cultural adjustment to America (5).

#### F. DISCUSSION

Almost half of the nonreligious subjects identified themselves as religious when living in America (Table 3). Greatest change was among the erstwhile Reform; little change occurred among the Orthodox. Actually there has been little change even among the Reform, for they can now with increased accuracy express their religious identification. In America it was practical to identify oneself as a nominal Jew (7). Nominal Jewish religious identification is unnecessary in Israel.

Americans who still identify themselves as religiously oriented also demonstrate greater Jewish identification while in America. Their parents' homes were religious and Zionist and they received a more intensive elementary Jewish education, followed by greater attendance at higher schools of Jewish learning (Table 2). Religiously oriented subjects also participated to a greater degree in Jewish community activities, more readily joining Zionist groups, synagogues, and Hillel on the college campus (Table 3), than did the nonreligious.

The religiously oriented were stimulated to visit Israel at an earlier age than the nonreligious. However, the religious took just as long in completing the act of settling. There was no group difference regarding the number of trips or years from initial trip before settling in Israel. The group with the more positive Jewish identification found it just as difficult to leave the security of America to begin life in a new land as did those not as strongly motivated to feel close to Israel.



The reactions of both groups concerning areas of life found to be disturbing in America suggest a difference in emphasis. The more positively identified religious oriented worry more about the group assimilating; whereas the nonreligious were more concerned about problems relating the individual to the total community, such as conformity and political witchhunting. The two groups' equal concern about anti-Semitism is a phenomenon also observed among Jewish adolescents (22) and college students (23).

## REFERENCES

1. BOGUE, D., & DIZARD, J. Race, ethnic prejudice, and discrimination as viewed by subordinate and superordinate groups. Mimeographed report, Community and Family Study Center, University of Chicago, Chicago, Illinois, April 16, 1964.
2. BOROWITZ, E. B. Hope Jewish and hope secular. *Judaism*, 1968, 17 (No. 2), 131-147.
3. CHENKIN, A. Jewish population in the United States, 1966. In *American Jewish Yearbook 1967* (Vol. 68). New York: American Jewish Committee, 1967. Pp. 231-242.
4. ENGEL, G. Some college students responses concerning Negroes of differing religious background. *J. Soc. Psychol.*, 1968, 74, 275-283.
5. GALLUP, G. Majority reject notion that U.S. is "sick society." American Institute of Public Opinion, Princeton, New Jersey, July 3, 1968.
6. GOLOVENSKY, D. I. The marginal man concept: An analysis and critique. *Soc. Forces*, 1952, 30, 333-339.
7. HERBERG, W. Protestant-Catholic-Jew. Garden City, N.Y.: Doubleday, 1955.
8. HERMAN, S. N. American Jewish students in Israel. *Jewish Soc. Stud.*, 1962, 24 (No. 1), 3-29.
9. ISAACS, H. R. American Jews in Israel. New York: Day, 1967.
10. KELMAN, H. C. The induction of action and attitude change. Proceedings of the XIV International Congress of Applied Psychology, Copenhagen, Denmark, 1961, 81-110.
11. ———. Discrepant action and attitude change: A functional analysis. Dittoed colloquium presented at University of Oregon, Eugene, April 28, 1967; and at University of California, Santa Barbara, May 16, 1967.
12. KING, B. T., & JANIS, I. L. Comparison of the effectiveness of improvised versus non-improvised role playing in producing opinion changes. *Hum. Relat.*, 1956, 9, 177-186.
13. LAPIDE, P. E. A Century of U.S. Aliya. Jerusalem, Israel: Jerusalem Post Press, 1961.
14. LEWIN, K. Psycho-sociological problems of a minority group. Chap. 9 in *Resolving Social Conflicts*. New York: Harper, 1948. Pp. 145-158.
15. MORRIS, Y. On the Soil of Israel. Tel Aviv, Israel: Association of Americans & Canadians in Israel, 1965.
16. PETUCHOWSKI, J. J. Zion Reconsidered. New York: Twayne, 1966.
17. ROSENTHAL, E. Acculturation without assimilation? The Jewish community of Chicago, Illinois. *Amer. J. Sociol.*, 1960, 66, 275-288.
18. ———. Studies of Jewish intermarriage in the United States. In *American Jewish Yearbook 1963* (Vol. 64). New York: American Jewish Committee, 1963. Pp. 3-53.

19. ———. Jewish intermarriage in Indiana. In *American Jewish Yearbook 1967 (Vol. 68)*. New York: American Jewish Committee, 1967. Pp. 243-264.
20. SANUA, V. C. Minority status among Jews and their psychological adjustment. *J. Jewish Sociol.*, 1962, 4 (No. 2), 242-253.
21. ———. Intermarriage and psychological adjustment. In *Marital Counseling: Psychology, Ideology, Science*, Silverman, H. L. (Ed.). Springfield, Ill.: Thomas, in preparation.
22. ———. Jewish education and attitudes of Jewish adolescents. *Educators Assembly Year Book of the United Synagogue of America*, 1967, 4, 1-26.
23. SARNOFF, I. Identification with the aggressor: Some personality correlates of anti-Semitism among Jews. *J. Personal.*, 1951, 20, 199-218.
24. SKLARE, M., & GREENBLUM, J. Jewish Identity on the Suburban Frontier. New York: Basic Books, 1967.
25. STONEQUIST, E. V. *The Marginal Man*. New York: Russell & Russell, 1937.
26. WEINGARTEN, M. *Life in a Kibbutz*. Jerusalem, Israel: Youth & Hehalutz Dept., 1959.

*B'nai B'rith Hillel Foundation at  
Purdue University  
State and Waldron  
West Lafayette, Indiana 47906*

## PROBLEMS IN THE CROSS-CULTURAL STUDY OF SCHIZOPHRENIA\*

*Department of Psychology, University of Calgary, Canada*

---

IHSAN AL-ISSA

---

### A. INTRODUCTION

In recent years there has been an increasing interest in the cross-cultural study of schizophrenia. The majority of these studies have been carried out by psychiatrists and anthropologists. Presently, there is a biannual review published in McGill University, which is devoted to the abstraction of studies in cross-cultural psychiatry. The present paper discusses some issues in the field of cross-cultural research that are relevant to schizophrenia in particular and mental disorder in general.

### B. GENERAL PROBLEMS

A general trend in the cross-cultural study of mental disorders is to deal with variations in their occurrence and symptomatology across cultures. In the early thirties it was claimed that schizophrenia is nonexistent in some non-Western cultures (10, 19); and this gave rise to some speculations about the etiology of schizophrenia, such as the complexity of Western culture (9). Later studies showing that the occurrence of schizophrenia is known in the West is almost universal (4, 32) have shifted the emphasis from the cross-cultural variations in occurrence to variations in symptomatology. Cultural variations in the occurrence and symptomatology of schizophrenia have been extensively reviewed by Mishler and Scotch (21) and by Al-Issa (1).

Among the most extensive cross-cultural studies of schizophrenic symptomatology is the one conducted by Murphy *et al.* (24). They sent a questionnaire to psychiatrists representing different cultures. Their findings showed some definite associations between the frequency of symptoms and certain cultural backgrounds. Delusions of destructiveness and religious delusions are frequently reported only among Christians and Muslims. Delusional jealousy is most frequent among Asians regardless of religion; visual hallucinations are reported most often in Africa and the Near East. Urban and rural pa-

---

\* Received in the Editorial Office, Provincetown, Massachusetts, on November 4, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

tients seem to differ in their symptoms. While the presence of depersonalization is seen most frequently among urban patients, delusions of grandeur seem to occur among rural patients. One important finding by Murphy *et al.* (24) is that such symptoms as social and emotional withdrawal, auditory hallucinations, general delusions, and flatness of affect appear to occur frequently in all their samples. As has been the case in the majority of cross-cultural studies of schizophrenic symptoms, the results may reflect the sociocultural background of the psychiatrist rather than true variations in symptomatology. Other problems involved in the use of the questionnaire or the survey method have been recently discussed by Wittkower and Rin (32) and by Al-Issa (1). However, the finding that certain symptoms are universal may have some implications for the study of abnormal behavior. It is often suggested by authors in the cross-cultural field that some symptoms cannot be accepted universally as signs of abnormality because of differences in the degree of their acceptance from one culture to the other: e.g., hallucinations involving religion are accepted in some cultures. This suggestion may be applicable to deviations where social evaluations are important in distinguishing between normality and abnormality (e.g., sexual deviation, nonconformity with the rules of society). Deviations involving cognitive symptoms which appear to be universal represent failure to adapt to the physical rather than the social environment. These cognitive abnormalities involve the inability to perceive correctly the nature of the physical world and do not depend on social evaluation. Whether a person is hallucinating or not is relevant to physical stimuli and can be studied objectively irrespective of cultural background. Reality in the sense of recognizing the physical world is a universal issue, and failure to be in contact with reality should be considered universally abnormal (6).

When cross-cultural studies of specific cognitive symptoms are examined, there appears to be a confusion between disturbances of certain functions and the indices of these functions. For instance, the frequent reports (17, 18, 20, 25) that paranoid delusions are very rare among illiterate peoples may reflect this confusion. It is possible that while paranoid delusions may exist among illiterate and literate peoples alike they tend to be expressed differently among these groups: e.g., expressed among illiterates in action (hitting or killing somebody) rather than in thought (during the clinical interview) (2). While disturbances in specific functions can be detected transculturally (misperception of reality), it is doubtful whether the same applies to the manifestations of these functions (content of hallucinations).

When the global concept of schizophrenia is used transculturally, the researcher is, of course, faced with increasingly complicated problems. Schizo-

phrenia is only a name for symptoms whose contents depend to a considerable extent on what a particular culture values or on what psychologists like to include within that concept. If the concept of schizophrenia is used for a non-Western person, one may end in dealing with something which has many different features when compared with its counterpart in the West. African schizophrenia, for instance, has been described as a poor imitation of European forms. In other words, the difference between the behavior patterns of these non-Western and Western patients is so great that one wonders whether it is legitimate to use the concept of schizophrenia at all. This observation is not surprising, since the reliability of this concept is doubtful even when used in the West (11). It is therefore meaningful to study specific functions and the effect of certain controllable variables on their development.

### C. THE AIM OF CROSS-CULTURAL RESEARCH

The study of specific functions of schizophrenic or other mental disorders may, of course, be carried out within a single culture, as well as between cultures. However, cross-cultural research would achieve two advantages over research designed within a single culture. First, for testing hypotheses, cross-cultural research provides homogeneous cultures that offer greater extremes on relevant variables than a single culture offers. In this case, the researcher makes use of the modal behavior of a reasonably homogeneous culture as the unit (in ordinary research the modal behavior of the individual rather than the whole society is taken as the unit). Irvin I. Child and his associates at Yale (3), for instance, explored the association of crime with child training practices, economy, and social structure in 48 preliterate societies. These cultural features show wider variations among these societies than within a single society, thus permitting a more comprehensive test of their significance. If one has an hypothesis about parent-child relationship and its effect on schizophrenic or other abnormal behavior, different societies, for instance, provide different degrees of contact with parents (23). You may compare societies with the following four types of households. (a) *The monogamous nuclear household*, which is the usual one in Western society. The father, mother, and children live under one roof. Grandparents, siblings of the parents, and other relatives live elsewhere. The effective presence of the father in the child's environment is thus at a maximum. (b) *The monogamous extended household* where two or more nuclear families live together under one roof. A typical extended family consists of an aged couple together with their married sons and daughters and their respective families. In such a household, the child's interaction with his father is likely to be somewhat less than in the



single nuclear household. (c) *Polygynous-polygynous household* which consists of a man living with his wives and their various children. Here the child is likely to have even less opportunity to interact with his father. Finally there is (d) the *polygynous mother-child household* where each wife has a separate establishment and lives in it with her children. In these societies the father either has a hut of his own, or divides his time among the houses of his various wives. The husband usually does not sleep in the house of any wife during the two or three years when she is nursing each infant. Thus the mother may become the almost exclusive object of contact with the child for the first few years of life (3).

Other differences in marriage rules such as endogamy (the habit of marrying within the tribe) may also provide an opportunity to check on the relative effect of hereditary and environmental factors on the incidence of pathological behavior.

If a society is taken as a single case in the study of a specific characteristic, our measure of that characteristic may represent the average among its many individuals. This measure is likely to be more reliable than a measure of the same tendency in a single individual or in a small segment of the society. It is certainly fruitful to carry out studies where the real incidence rather than the diagnosed prevalence of certain behavior pathology is investigated. This can be achieved in small and circumscribed societies.

With Westernization, societies are starting to change rapidly and the author feels that this is the end of the Golden Age for cross-cultural research before the evergrowing contact between cultures has lessened the differences between societies.

A second advantage of cross-cultural research is to provide appropriate conditions for the systematic variation of factors that cannot be varied within a single culture. If some cultural characteristic is believed to have interacting effects on the variables between which relational principles are being sought, then testing of the hypotheses must be done in different cultures that can provide such interaction (28). An interesting example is the recent work in Africa on factors affecting visual illusions. There are, for instance, the effects of the ability to interpret two-dimensional representation and the degree of rectangularity in the environment on the Müller-Lyer illusion (13, 15). Let us suppose that we are interested in the interaction between literacy and ecology on perception. We may find that the difference in the perception of literates and illiterates varies in these groups according to their living surroundings. Some of the possible interactions between literacy and ecology on visual illusion in two hypothetical cultures may be as shown in Table 1.

In small tribal societies it should be possible to collect data on role conflict and interpersonal influence that are difficult to isolate in a society with a multiplicity of occupations, a high level of literacy, and well developed mass media of communication. In an illiterate society where the only channels of communication are through face to face contact and the system of social situations is fairly rudimentary, it seems feasible to press beyond gross variables (social class) to specific features of the sociocultural environment (16). Stressful situations and their effect on the incidence of mental disorders could be more

TABLE 1  
SOME POSSIBLE INTERACTIONS BETWEEN LITERACY AND ECOLOGY ON  
VISUAL ILLUSION IN TWO HYPOTHETICAL CULTURES

Interactions	Culture I (carpentered environment)	Culture II (noncarpentered environment)
	Literates vs. illiterates	Literates vs. illiterates
I	No difference	Literates > Illiterates
II	No difference	Illiterates > Literates
III	Literates > Illiterates	Illiterates > Literates
IV	Literates > Illiterates	Literates > Illiterates*

\* The difference is significantly larger than in Culture 1.

easily delineated in these societies than in literate societies. We may obtain interesting information by comparing the incidence of mental or psychosomatic disorders among a group of first wives with a group of "favored" second wives or a group of the present rulers with another group of deposed rulers as a result of a local *coup d'état*.

#### D. THE PROBLEM OF CULTURE-FREE MEASUREMENT

Whether the culture is used as a homogeneous unit or a source of systematic variation of an interaction variable, the investigator should search for variables that are measurable in whatever culture is chosen. Early cross-cultural studies gave considerable attention to the construction and use of culture-free tests, but these attempts have not been very successful. When projective techniques, which were regarded as culture-free, were applied to non-Western peoples, the test records regularly showed signs of psychopathy as if everybody in those communities were sick but the tester.

Two early studies by Bleuler and Bleuler (5) and Cook (8), which used the Rorschach test, may illustrate this point. The desert Moroccans give responses to hardly observable details, while Samoans give whole responses to the entire blot with a tendency to respond to the white space. According to Western criteria, the responses of the Moroccans and the Samoans would

warrant the diagnosis of schizophrenia (31, p. 314). Of course, such results are not surprising, since these tests have doubtful validity even when used in the West.

It should be noted here that the problem of the degree to which a test is culture-free should not be confused with the problem of whether or not a test is culturally appropriate for the measurement of the variables under investigation. A test is appropriate if the test items are familiar to the cultural group. Regardless of the interaction of different cultural variables on the response under investigation, the familiarity of subjects with the test material may affect the results. In his investigation of abstract abilities of Africans, Jahoda (14) found significant differences between his literate and illiterate subjects on the Goldstein Scheerer Cube Test. These results appear to support the assumption that illiterate Africans are characterized by concreteness and lack of abstraction which is similar to Western brain-damaged or schizophrenic subjects (7, 12). Using familiar material taken from the children's environment (plant and animal material), Price-Williams (26) found no difference between illiterates and literates. Furthermore, he found that both literate and illiterate children were able to reach the various stages of performance in classification tasks if the same familiar material was used.

Tests may differ in their appropriateness to a certain culture, as well as in the degree to which they are influenced by cultural differences. While the choice of appropriate test items may be possible, it is doubtful that there are tests which are completely culture-free. Therefore, instead of looking for culture-free tests, one should use tests which, though amenable to the influence of cultural forces, still permit the measurement of the functions investigated. In this case, choice of tests would be based on the degree to which culture affects different measures. It may be possible to arrange responses according to the degree to which they are influenced by culture. Conceptual responses may be regarded as the most amenable to cultural influence; while physiological responses as the least, with psychomotor, perceptual, and sensory responses falling in between. It should be noted that *the way* in which culture influences these categories of responses is also different. Zubin (33) described these influences as follows: "Though no measures can be said to be completely culture-free, the way in which culture affects certain measures (as pupillary response to light stimuli) is indirect, unlike the direct way in which it influences primarily conceptual measures like vocabulary. The major way in which culture will tend to influence the culture-free or culture-fair tests is probably not in the function under measurement, but in the subject's approach to the testing situation: e.g., in the subject's understanding of the purpose of the test, in the

degree of fear, in his motivation, attention, and cooperation, etc. In other words, the influence of culture is specifically on those variables which also tend to contaminate comparisons of schizophrenics and normals even when they come from the same cultural background" (33, pp. 66-67). The findings of Robertson and Batchelder (27) are relevant to the "indirect" influences of culture in the test situation. They found that British subjects tended to stress accuracy at the expense of speed when compared with the American norms of the Wechsler Adult Intelligence Scale. However, these influences are controllable in the test situation, such as by emphasizing speed or accuracy in the instructions.

Cross-cultural research should start with the measurement of these responses that are minimally influenced by culture. The patterns of these responses should be ascertained among normal subjects cross-culturally. The study by Mundy-Castle *et al.* (22) comparing the EEG of normal Africans and Whites illustrates this approach. The finding that there is no significant difference in the EEG patterns of these two groups may favor the use of this measure cross-culturally. It is perhaps useful to go a step further by also comparing the EEG or other physiological responses of abnormal subjects across cultures. Since these responses are minimally influenced by culture, results showing similarities between abnormals with different cultural backgrounds may give us some indications of organic factors in the etiology of behavior pathology. Of course, we are again faced with the problem of finding a criterion for normality and abnormality that can be applied cross-culturally. We have pointed out in the first section of this paper that certain symptoms (e.g., the cognitive symptoms of hallucination and delusion) tend to occur in different cultures; and these symptoms may be taken as our criterion for differentiating between normals and abnormals across cultures.

The attempt to establish normative data of physiological responses of normals and abnormals within and across cultures may lead us to compare these responses of abnormal subjects with subjects who show the same response pattern but without behavioral pathology. Though studies concerning physiological responses, which are minimally influenced by culture, are lacking, we may illustrate our point from the field of perception. Using Indian and British students, Thouless (29) found that Indians show greater tendency toward object constancy than his British subjects. Recent experimental work shows that Western schizophrenics tend to show underconstancy (30). However, paranoid schizophrenics show overconstancy, a result similar to that reported by Thouless concerning Indian students. Here, non-Western normals are found to respond in a similar way to paranoid schizophrenic subjects in the West.



This research scheme which is advocated by Professor Zubin at Columbia University (33) may give us some indications of the types of interactions between perceptual (or physiological) and sociocultural variables which led to schizophrenia in one group (e.g., British or American) but not in the other (e.g., Indian).

In brief, the issues outlined in this paper are common to psychologists working within, as well as across, cultures. Cross-cultural research only widens the scope and highlights the difficulties within which the psychologist works in his search for factors affecting normal and abnormal behavior.

### E. SUMMARY

The recent interest in the cross-cultural study of mental disorders has brought about many research problems in this field of study. These problems have been extensively discussed by Mishler and Scotch (21), Al-Issa (1), and Wittkower and Rin (32). The present paper has dealt with the specific problem of validity of measures used across cultures. It is suggested that physiological responses are minimally influenced by culture. These responses may be used as a baseline in comparing normal and schizophrenic subjects within, as well as across, cultures.

### REFERENCES

1. AL-ISSA, I. Cross-cultural studies of symptomatology in schizophrenia. *Can. Psychiat. Assoc. J.*, 1968, **13**, 147-157.
2. ———. Literacy and symptomatology in chronic schizophrenia. *Brit. J. Soc. Psychiat.*, 1967, **1**, ———.
3. BACON, M. K., CHILD, I. L., & BARRY, H. A cross-cultural study of correlates of crime. *J. Abn. & Soc. Psychol.*, 1963, **66**, 291-300.
4. BENEDICT, P. K., & JACKS, I. Mental illness in primitive societies. *Psychiatry*, 1954, **17**, 377-389.
5. BLEULER, M., & BLEULER, R. Rorschach ink-blot tests and social psychology. *Charac. Personal.*, 1934, **4**, 99-114.
6. BUSS, A. Psychopathology. New York: Wiley, 1966.
7. CAROTHERS, J. C. The African mind in health and disease. Geneva, Switzerland: World Health Organization, 1953.
8. COOK, T. H. The application of the Rorschach test to a Samoan group. *Rorschach Res. Exch.*, 1942, **6**, 51-60.
9. DEVEREUX, G. A sociological theory of schizophrenia. *Psychoanal. Rev.*, 1934, **26**, 315-342.
10. DHUNJIBOY, J. Brief résumé of the types of insanity commonly met with in India. *J. Ment. Sci.*, 1930, **16**, 254-264.
11. GUERTIN, W. H. Medical and statistical-psychological models for research in schizophrenia. *Behav. Sci.*, 1961, **6**, 200-204.
12. HAWARD, L. C. R., & ROLAND, W. A. Some inter-cultural differences on the Draw-a-Man Test: Goodenough scores. *Man*, 1954, **54**, 86-88.



13. HUDSON, W. Pictorial depth perception in sub-cultural groups in Africa. *J. Soc. Psychol.*, 1960, 52, 183-208.
14. JAHODA, G. Assessment of abstract behavior in a non-Western culture. *J. Abn. & Soc. Psychol.*, 1956, 53, 237-243.
15. ———. Geometric illusions and environment: A study in Ghana. *Brit. J. Psychol.*, 1966, 57, 193-199.
16. KENNEDY, D. A. Key issues in the cross-cultural study of mental disorders. In B. Kaplan (Ed.), *Studying Personality Cross-Culturally*. Evanston, Ill.: Row, Peterson, 1961.
17. KLINE, N. S. Psychiatry in Indonesia. *Amer. J. Psychiat.*, 1963, 119, 809-815.
18. ———. Psychiatry in Kuwait. *Brit. J. Psychiat.*, 1963, 109, 766-774.
19. LOPEZ, C. Ethnographische Betrachtungen über schizophrenie. *Ztschr. Ges. Neurol. u. Psychiat.*, 1932, 142, 706-711. Cited in M. K. Opler (Ed.), *Culture, Psychiatry and Human Values*. Springfield, Ill.: Thomas, 1956.
20. LOUDON, J. B. Psychogenic disorder and social conflict among the Zulu. In M. K. Opler (Ed.), *Culture and Mental Health*. New York: Macmillan, 1959.
21. MISHLER, E. G., & SCOTCH, N. A. Sociocultural factors in the epidemiology of schizophrenia. *Psychiatry*, 1963, 26, 315-351.
22. MUNDY-CASTLE, A. C., MCKIEVER, B. L., & PRINSLOO, T. A comparative study of the electroencephalograms of normal Africans and Europeans of Southern Africa. *EEG Clin. Neurophysiol.*, 1953, 5, 533-543.
23. MURDOCK, G. P. World ethnographic sample. *Amer. Anthropologist*, 1957, 59, 664-687.
24. MURPHY, H. B. M., WITTKOWER, E. D., FRIED, J., & ELLENBERGER, H. A cross-cultural survey of schizophrenic symptomatology. In *Proceedings of the Third World Congress of Psychiatry*, 1961, 2, 1309-1315 (Montreal, Canada: McGill Univ. Press).
25. OPLER, M. K. Culture psychiatry and human values. Springfield, Ill.: Thomas, 1956.
26. PRICE-WILLIAMS, D. R. Abstract and concrete modes of classification in a primitive society. *Brit. J. Educ. Psychol.*, 1962, 32, 50-61.
27. ROBERTSON, J. P. S., & BATCHELDOR, K. J. Cultural aspects of the Wechsler Adult Intelligence Scale in relation to British mental patients. *J. Ment. Sci.*, 1956, 102, 612-618.
28. SEARS, R. R. Transcultural variables and conceptual equivalence. In B. Kaplan (Ed.), *Studying Personality Cross-Culturally*. Evanston, Ill.: Row, Peterson, 1961.
29. THOULESS, R. H. A racial difference in perception. *J. Soc. Psychol.*, 1933, 4, 330-339.
30. VENABLES, P. H. Input dysfunction in schizophrenia. In B. A. Maher (Ed.), *Progress in Experimental Personality Research (Vol. 1)*. New York: Academic Press, 1964.
31. WEINER, I. B. Psychodiagnosis in Schizophrenia. New York: Wiley, 1966.
32. WITTKOWER, E. D., & RIN, H. Transcultural psychiatry. *Arch. Gen. Psychiat.*, 1965, 13, 387-394.
33. ZUBIN, J. A cross-cultural approach to psychopathology and its implications for diagnostic classification. In L. D. Eron (Ed.), *The Classification of Behavior Disorders*. Chicago, Ill.: Aldine, 1966.

Department of Psychology  
The University of Calgary  
Calgary, Alberta, Canada



# Journal of Genetic Psychology

Founded by G. Stanley Hall in 1891; reorganized by Carl Murchison in 1925

Quarterly, beginning with March. Devoted to child behavior, animal behavior, and the problems of aging. Manuscripts may be sent to any member of the Editorial Board or directly to The Journal Press. Subscriptions should be sent directly to The Journal Press, and business communications addressed to the Managing Editor. Two volumes annually. Per annum, \$20.00; per volume, \$15.00; single numbers, \$7.50. Complete sets from 1891 at \$15.00 per volume, plus transportation.

---

# Journal of General Psychology

Founded in 1927 by E. B. Titchener and Carl Murchison

Quarterly, beginning with January. Devoted primarily to experimental, theoretical, clinical, and historical psychology. Manuscripts may be sent to any member of the Editorial Board or directly to The Journal Press. Subscriptions should be sent directly to The Journal Press, and business communications addressed to the Managing Editor. Two volumes annually. Per annum, \$20.00; per volume, \$15.00; single numbers, \$7.50. Complete sets from 1928 at \$15.00 per volume, plus transportation.

---

# Journal of Social Psychology

Founded in 1929 by John Dewey and Carl Murchison

Bimonthly, beginning with February. Devoted to studies of persons in group settings and of culture and personality; special attention to cross-cultural articles and notes, and to briefly reported replications and refinements. Manuscripts may be sent to any member of the Editorial Board or directly to The Journal Press. Subscriptions should be sent directly to The Journal Press, and business communications addressed to the Managing Editor. Three volumes annually. Per annum, \$30.00; per volume, \$15.00; single numbers, \$7.50. Complete sets from 1930 at \$15.00 per volume, plus transportation.

---

*Each author will receive 100 unbound reprints gratis. Additional reprints may be purchased at a nominal cost when the author returns page-proof.*

THE JOURNAL PRESS  
2 Commercial Street  
Provincetown, Massachusetts 02657  
U. S. A.



# Journal of Genetic Psychology

Founded by G. Stanley Hall in 1891; reorganized by Carl Murchison in 1925

Quarterly, beginning with March. Devoted to child behavior, animal behavior, and the problems of aging. Manuscripts may be sent to any member of the Editorial Board or directly to The Journal Press. Subscriptions should be sent directly to The Journal Press, and business communications addressed to the Managing Editor. Two volumes annually. Per annum, \$20.00; per volume, \$15.00; single numbers, \$7.50. Complete sets from 1891 at \$15.00 per volume, plus transportation.

---

# Journal of General Psychology

Founded in 1927 by E. B. Titchener and Carl Murchison

Quarterly, beginning with January. Devoted primarily to experimental, theoretical, clinical, and historical psychology. Manuscripts may be sent to any member of the Editorial Board or directly to The Journal Press. Subscriptions should be sent directly to The Journal Press, and business communications addressed to the Managing Editor. Two volumes annually. Per annum, \$20.00; per volume, \$15.00; single numbers, \$7.50. Complete sets from 1928 at \$15.00 per volume, plus transportation.

---

# Journal of Social Psychology

Founded in 1929 by John Dewey and Carl Murchison

Bimonthly, beginning with February. Devoted to studies of persons in group settings and of culture and personality; special attention to cross-cultural articles and notes, and to briefly reported replications and refinements. Manuscripts may be sent to any member of the Editorial Board or directly to The Journal Press. Subscriptions should be sent directly to The Journal Press, and business communications addressed to the Managing Editor. Three volumes annually. Per annum, \$30.00; per volume, \$15.00; single numbers, \$7.50. Complete sets from 1930 at \$15.00 per volume, plus transportation.

---

*Each author will receive 100 unbound reprints gratis. Additional reprints may be purchased at a nominal cost when the author returns page-proof.*

THE JOURNAL PRESS  
2 Commercial Street  
Provincetown, Massachusetts 02657  
U. S. A.



# Club Rate for Five International Psychological Journals

	PER ANNUM
THE JOURNAL OF GENETIC PSYCHOLOGY . . . . .	\$20.00
(Quarterly, two volumes per year)	
GENETIC PSYCHOLOGY MONOGRAPHS . . . . .	\$20.00
(Quarterly, two volumes per year)	
THE JOURNAL OF GENERAL PSYCHOLOGY . . . . .	\$20.00
(Quarterly, two volumes per year)	
THE JOURNAL OF SOCIAL PSYCHOLOGY . . . . .	\$30.00
(Bimonthly, three volumes per year)	
THE JOURNAL OF PSYCHOLOGY . . . . .	\$30.00
(Bimonthly, three volumes per year)	
	TOTAL \$120.00
TEACHERS AND STUDENTS EDITION, <i>Restricted Sale</i>	\$50.00

The *Teachers and Students Edition* is open to any individual psychologist, teacher, student, or physician but requires an agreement that the edition will not be offered for sale till after five years from publication and an agreement that the edition will not be substituted for an institutional subscription.

THE JOURNAL PRESS  
2 Commercial Street  
Provincetown, Massachusetts 02657  
U. S. A.

## Genetic Psychology Monographs

Founded in 1925 by Carl Murchison

Quarterly, beginning with February. Devoted to child behavior, animal behavior, and the problems of aging. Each number one or more complete researches. Manuscripts may be sent to any member of the Editorial Board or directly to The Journal Press. Subscriptions should be sent directly to The Journal Press, and business communications addressed to the Managing Editor. Two volumes annually. Per annum, \$20.00; per volume, \$15.00; single numbers, \$7.50. Complete sets from 1926 at \$15.00 per volume, plus transportation.

**NOTICE.** Effective immediately we offer prior publication for certain selected monographs. Charges will be based on the announced rates for *The Journal of Psychology*. For a limited period of time our policy is changed for regular monographs: Authors will be charged for cuts and expensive tables, but no page charge will be imposed. Two hundred reprints of each monograph will be supplied gratis.

The listing of monographs for 1926 through 1963 will be printed in the first half of each volume of all five journals. Monographs published after 1963 will be listed in the second half of each volume.

# Genetic Psychology Monographs (continued)

VOLUME 69—January-June, 1964

- Occupational and industrial therapy in the treatment of psychiatric patients: A controlled study of efficacy in an intensive-treatment institution—M. HUNDZIAK AND B. PARAMANICK  
An analysis of Carl R. Rogers' theory of personality—M. S. KRAUSE  
Personality aspects of teaching: A predictive study—H. ZIMMEL, B. BIKER, W. RADINOWITZ, AND L. HAY  
Situational and personal antecedents of incipient alienation: An exploratory study—W. H. JARRETT AND A. O. HALLER
- Age changes in children's mosaic responses from five to ten years—L. B. AMES AND F. L. ILC  
A study of the interview process: The caseworker-client relationship—H. J. MEYER, E. F. BORGATTA, AND D. FANSHILL  
A study of introversion-extroversion as a personality dimension—R. M. ALLEN, H. M. RICHNER, AND R. J. PLOTNIK  
Psychological testing with Northwest Coast Alaskan Eskimos—C. E. PRESTON  
A study in family pathology—M. B. HANS

VOLUME 70—July-December, 1964

- Structure and function—H. GUTMAN  
The Lowenfeld Mosaic Test: Norms for five-to-ten-year-old American public-school children and comparative study of three groups—L. B. AMES, F. L. ILC, AND J. AUGUST  
Social and personality correlates of children's estimates of height—J. P. SHAFFER  
The development of ethnic attitudes in New Zealand school children—G. M. VAUGHAN
- The antecedents of two varieties of vocational interests—L. E. TYLER  
Personality in old age—P. E. SLATER AND H. A. SCARR  
Variables affecting experimenter bias in a group situation—R. ROSENTHAL, C. J. FRIEDMAN, C. A. JOHNSON, K. FODE, T. SCHILL, R. C. WHITE, AND L. L. VIKAN  
Ordinal position and sex role identification—B. G. ROSENBERG AND B. SUTTON-SMITH  
Personality correlates of accuracy of self-appraisal in four-year-old children—PAUL S. WEINER

VOLUME 71—January-June, 1965

- Author index for Genetic Psychology Monographs 1926-1964—L. W. FERGUSON  
A theoretical, developmental, and clinical investigation of embarrassment—J. M. SATTLER  
School readiness as evaluated by Gesell developmental, visual, and projective tests—F. L. ILC, L. B. AMES, AND R. I. APPELL  
Rorschach scales of regression in psychosis—S. NECKER, G. ALAN, AND D. A. QUIRK  
The effects of extra tactile stimulation on a group of institutionalized infants—L. CASLER  
Human eyelid conditioning: The recent experimental literature—L. E. ROSS AND T. F. HARTMAN
- Social status and schoolroom aggression in third-grade children—R. TOICO  
Overt and covert homosexual problems among male college students—L. J. BRAATEN AND C. D. DARLING  
Role conceptions of the psychiatric aide—M. P. LAWTON AND A. E. GOLDMAN  
Changes in study of values scores during the four years of college—C. W. HUNTLEY

VOLUME 72—July-December, 1965

- A study of process and method in three-year-old twins and triplets learning to read—W. FOWLER  
Individualized teaching with therapeutic aims: A methodological study—M. KOHN AND M. RUDNICK  
Ability and social adjustment at midlife of persons earlier judged mentally deficient—E. L. MILLER
- Development of incentive values in childhood—S. L. WITKOW, D. J. TYRRELL, AND L. M. LOWDEN  
Longitudinal changes in intolerance and authoritarianism for subjects differing in amount of college education over four years—W. T. PLANT  
An NDEA counseling and guidance institute: Prediction, performance, and follow-up—R. H. DANA  
The influence of language on classification: A theoretical model applied to normal, retarded, and deaf children—H. G. FURTH AND N. A. MILGRAM  
Time estimates as affected by incentive class and motivational level—M. FELIX  
Clinical psychology: An historical survey—V. S. SEXTON

VOLUME 73—January-June, 1966

- Vulnerability in infants and young children: A pilot study—G. HEIDIK
- Racial preference and ego-identity among white and Bantu children in the republic of South Africa—A. J. GREGOR AND D. A. MCPHERSON  
Mother-child interaction in cerebral palsy: Environmental and psychosocial obstacles to cognitive development—E. SHERR AND R. KASTENBAUM  
Children's stories—L. B. AMES

VOLUME 74—July-December, 1966

- Changes in personality for groups completing different amounts of college over two years—W. T. PLANT AND C. W. TELFORD  
Brain dynamism in perception of reversible figures—G. K. YACORZYNSKI  
Changes in Rorschach response throughout the human life span—L. B. AMES  
Mnemonic virtuosity: A study of chess players—A. BIKER (translated by M. L. SIMMEL AND S. B. BAYTON)
- The development of behavior patterns in children—B. B. TYLER, F. B. TYLER AND J. E. RAFFERTY  
Correlates of classroom functioning—L. M. SMITH AND B. B. HUDGINS  
Children's behavior syndromes and parental responses—R. L. JENKINS, E. NUREDDIN, AND I. SHAPIRO

VOLUME 75—January-June, 1967

- Studies of an elementary game of strategy—B. SUTTON-SMITH AND J. M. ROBERTS  
Child care practices anteceding three patterns of preschool behavior—D. BAUMRING  
The impact of cognitive maturity on the development of sex-role attitudes in the years 4 to 8—L. KOHLBERG AND E. ZIGLER
- Differential factor structures in mathematical ability—P. S. VERY  
Perception of the vertical as a function of affective stimuli and extraneous stimulation in schizophrenic and normal subjects—J. O. WELLS AND W. E. CALDWELL  
Mid-life attainment of the mentally retarded: A longitudinal study—W. R. BALLER, D. C. CHARLES, AND E. L. MILLER

VOLUME 76—July-December, 1967

- Scalogram analysis of some developmental sequences of infant behavior as measured by the Bayley Infant Scale of Mental Development—R. KOHEN-RAZ  
The leader and the lost: A case study of indigenous leadership in a poverty program Community Action Committee—L. A. ZURCHER  
A developmental study of the perception of people—R. N. WOLMAN  
Human fears: A factor analytic study of three age levels—G. W. RUSSELL
- Correlates of voluntary social participation—J. E. TIKLE  
Personality characteristics and developmental history of creative college women—R. HELSON  
A systematic study of therapeutic interventions in nursing homes—B. A. STOSKY

# Genetic Psychology Monographs (continued)

## VOLUME 77—January-June, 1968

1. The fulfillment of promise: 40-year follow-up of the Terman Gifted Group—M. H. ODEN  
A scale for parents' ratings: Some ipsative and normative correlations—R. N. WALKER  
A methodological investigation of Piaget's concept of conservation of substance.—A. B. BITTNER AND M. M. SHINEDLING
2. Sex differences in mental and behavioral traits—J. E. GARAI AND A. SCHEINFELD

## VOLUME 78—July-December, 1968

1. A theoretical and developmental investigation of delayed speech feedback—J. E. WATERS  
Growth changes in factorial structure by age and sex—N. W. DYE AND P. S. VERY  
A multispect investigation of masculine development in kindergarten age boys—H. B. BILLER
2. The optimum limits of testing on the Wechsler Intelligence Scales—M. Y. QUEKRESHI  
A developmental study of the ego-ideal—L. VAN DEN DAEL  
The relationship between physical condition immediately after birth and mental and motor performance at age four—N. EDWARDS

## VOLUME 79—January-June, 1969

1. Adolescents: From the views of the members of an informal adolescent group—T. P. JOSEPH  
Acculturation and childhood accidents among Chinese and Japanese Americans—M. KUROKAWA



## PUBLICATION IN THE JOURNAL OF PSYCHOLOGY

*The Journal of Psychology* provides a unique service to authors. Articles accepted for publication in this periodical appear in print with the least possible amount of delay. In line with this policy, all steps of the acceptance procedure are maximally streamlined, and the author's work is presented with a bare minimum of editing or restriction of any kind. *The Journal of Psychology* is the ideal medium for presentation of timely professional material.

Authors are requested to exercise unusual care in the preparation of papers for this journal. Accuracy, grammar, and style should be checked and rechecked before a clean version of the manuscript is submitted. Careful attention should be paid to the General Instructions on the back cover of each issue and to the announcement on the inside front cover.

Authors are particularly requested to note the fact that all parts of the manuscript—including references, footnotes, and tables—must be double-spaced.



\$30.00 per annum  
\$15.00 per volume  
\$7.50 per half volume

BIMONTHLY  
Three volumes per year  
Immediate publication

March, 1969  
Volume 71, Second Half

(Founded in 1935 by Carl Murchison)

# The Journal of PSYCHOLOGY

The General Field of Psychology

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold

under the restrictions of a bilateral contract, and may not be resold for a period of five years from the date of publication.

MARCH, 1969

(Manuscripts are printed in the order of final acceptance)

- Faking and faking detection on the 16 PF-Form A . . . . . 155  
By JOHN R. BRAUN AND DOLORES LAFARO
- The use of analogies and incubation in obtaining insights in creative problem solving . . . . . 159  
By ROY DREISTADT
- The effects of punishment on food pellet discrimination . . . . . 177  
By WILLIAM F. VITULLI
- Visual acuity decrement for a subfusional intermittency with two directions of border contrast . . . . . 185  
By S. HOWARD BARTLEY AND RICHARD J. BALL
- Problem-solving motivation in verbal conditioning studies . . . . . 191  
By ROGER E. VOGLER AND RUTH L. AULT

(OVER)

Copyright © 1969 by The Journal Press  
Provincetown, Massachusetts, U. S. A.

All rights reserved

Published at 35 New St., Worcester, Mass.

Second-class postage paid at Worcester, Mass.

Send all mail to Managing Editor, 2 Commercial Street, Provincetown, Massachusetts 02657

Factorial structure of CQT, ACT, and SAT test scores for 50 available college freshmen . . . . .	199
By RUSSELL N. CASSEL AND ARDEN C. EICHSTEADT	
Cognitive complexity, strategy of the other player, and two-person game behavior . . . . .	205
By JOSEPH G. PHELAN AND ED. RICHARDSON	
Relationships between expressed personality needs and social background and military status variables . . . . .	217
By MAXINE MURPHY GUNDERSON	
Evaluation of self and others and some personality correlates . . . . .	225
By DARHL M. PEDERSEN	
Effects of vicarious verbal stimuli on conditioning of hostile and neutral verbs . . . . .	245
By HOYT MELVYN GILLEY	
The effects of item difficulty sequencing and anxiety reaction type on aptitude test performance . . . . .	253
By VINCENT F. BERGER, DAVID C. MUNZ, ALBERT D. SMOUSE, AND HENRY ANGELINO	
A methodological note on "a self-concept inventory" . . . . .	259
By SHIRIN SCHLUDERMANN AND EDUARD SCHLUDERMANN	
Factorial study and analysis of factor variance of intellective self-ratings from a combined sample of male and female subjects . . . . .	261
By RIO SCIORTINO	
Factorial study of general adaptability self-ratings by male and female subjects . . . . .	271
By RIO SCIORTINO	
Brightness matching under intermittent stimulation . . . . .	281
By JONATHAN D. KAZSUK AND S. HOWARD BARTLEY	

## FAKING AND FAKING DETECTION ON THE 16 PF-FORM A\*<sup>1</sup>

*Department of Psychology, University of Bridgeport*

---

JOHN R. BRAUN AND DOLORES LAFARO

---

### A. INTRODUCTION

The 16 PF Questionnaire (16 PF) provides scores on 16 factor-analytically derived dimensions (3). Along with most other inventories, the 16 PF falls short of meeting the recommendation of the 1966 *Standards for Educational and Psychological Tests and Manuals* that "... evidence should be presented on the extent to which scores are susceptible to an attempt by the examinee to present a false or unduly favorable picture of himself" (1, p. 24).

Braun and LaFaro (2) have filled part of this gap in 16 PF data by studying the effects of faking instructions on Form C. They found that scores on eight of the 16 factors were significantly affected, and that the Form C Motivational Distortion (MD) scale was "... of quite limited usefulness in discriminating between control and faked protocols" (2, p. 7).

Unlike Form C, Form A does not routinely provide an MD score. Instead, the manual for Forms A and B (4, p. 5) stresses that items have been selected "... to be as neutral in value as possible" and "... to emphasize both desirable and undesirable aspects at both ends of each factor scale." Further, items which "... do not obviously refer to the trait but which correlationally are known to measure it, have been chosen wherever possible, as a 'built-in' protection against distortion." Since no evidence is presented to demonstrate the effectiveness of these item selection attempts to reduce faking problems, the present study investigated the effects of faking instructions on Form A scores. In addition, the usefulness of an experimental MD scale devised by Schanberger (7) was studied. Schanberger has found 15 Form A items that significantly differentiated between the responses of 88 nonpsychotic hospitalized offenders under "fake-good" versus standard instructions. We investigated the generality of this scale's effectiveness by determining whether it differentiated between responses under "fake-good" versus standard instructions when the Ss were university students warned to try to conceal their faking.

---

\* Received in the Editorial Office, Provincetown, Massachusetts, on November 6, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This report is based on a paper read by the junior author at the annual meeting of the Eastern Psychological Association, Washington, D.C., April, 1968.

## B. METHOD

Ss were undergraduates recruited from introductory psychology courses. Group I ( $N = 60$ ) had standard administration first, and one week later, the faked-administration. Group II ( $N = 41$ ) had only a single administration which was the faked one. Group I Ss used self-selected pseudonyms on their papers, while Group II Ss used their actual names.

The instructions for the faked administrations were identical with those used in the Form C study (2):

"One of the problems with personality questionnaires is that people may try to slant their answers so as to give a good impression of themselves. That is, instead of giving answers that reflect the way they actually feel or act, they give answers that they believe will place them in a good light. They may try to fake their answers somewhat.

"Personality inventories vary in the extent to which such deliberate faking is possible. With some it is very easy; with others quite difficult. We are studying how the 16 PF is in this regard. Contrary to the instructions given on the cover of the booklet, we want you to answer not necessarily with your true feelings, but rather in the way that you would if you were actually going to try to fake—to beat the test.

"However, we must warn you that the test has certain features by which it attempts to detect faking. Answer the test so that you make the best impression you can, but be careful to do it in such a way that your faking cannot be detected."

## C. RESULTS AND DISCUSSION

Means, standard deviations, and results of significance tests—Sandler's A-statistic (6, p. 172)—are presented in Table 1. Differences between standard and faked-administration means were significant for 11 of the 16 test factors. The direction of these significant differences was such that Ss, when faking, appeared outgoing rather than reserved, stable rather than affected by feelings, happy-go-lucky rather than sober, conscientious rather than expedient, venturesome rather than shy, trusting rather than suspicious, shrewd rather than forthright, placid rather than apprehensive, group-dependent rather than self-sufficient, controlled rather than undisciplined, and relaxed rather than tense.

Seven of the eight factors that were significantly influenced by faking instructions in the Form C study (2) were significantly changed in the same direction in the present investigation. The eighth of these factors ( $N$ —forth-

right *versus* shrewd) was significantly decreased on Form C, but significantly increased on Form A.

Although the present authors did not learn of it until quite recently, Handler and Shrader (5) have also investigated faking on the 16 PF, using three groups of Ss with instructions to look as well adjusted as possible, and without any cautionary directive about attempting to conceal faking. There is excellent agreement between the present findings and those of Handler and Shrader with regard to the direction and relative magnitude of changes in scores produced by faking instructions on the various 16 PF scales.

TABLE 1  
MEANS AND STANDARD DEVIATIONS FOR STANDARD AND FAKED ADMINISTRATIONS ( $N = 60$ )

Variable	Standard		Faked	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Factor A—reserved <i>versus</i> outgoing	10.77	3.14	13.53**	3.03
Factor B—dull <i>versus</i> bright	7.98	1.60	7.98	1.81
Factor C—affected by feelings <i>versus</i> stable	13.87	3.11	19.13**	4.33
Factor E—humble <i>versus</i> assertive	13.37	3.95	12.85	4.20
Factor F—sober <i>versus</i> happy-go-lucky	15.93	5.03	17.72*	3.88
Factor G—expedient <i>versus</i> conscientious	12.30	3.55	15.73**	3.83
Factor H—shy <i>versus</i> venturesome	12.03	5.03	19.25**	4.49
Factor I—tough-minded <i>versus</i> tender-minded	10.52	2.97	10.23	4.19
Factor L—trusting <i>versus</i> suspicious	9.45	2.89	7.23**	3.16
Factor M—practical <i>versus</i> imaginative	13.07	3.09	12.63	2.75
Factor N—forthright <i>versus</i> shrewd	10.55	2.78	13.10**	2.75
Factor O—placid <i>versus</i> apprehensive	11.48	3.44	6.83**	3.76
Factor Q <sub>1</sub> —conservative <i>versus</i> experimenting	10.32	2.60	10.33	3.45
Factor Q <sub>2</sub> —group-dependent <i>versus</i> self-sufficient	10.52	3.19	6.93**	2.82
Factor Q <sub>3</sub> —undisciplined <i>versus</i> controlled	9.13	2.84	13.53**	3.70
Factor Q <sub>4</sub> —relaxed <i>versus</i> tense	13.42	4.06	6.27**	5.48
Schanberger Motivational Distortion	3.62	1.91	9.23**	3.84

Note: For the factor scores, low scores are associated with the term listed first (e.g., Factor A—reserved); high scores with the second term (e.g., Factor A—outgoing).

\* Significantly different from Standard Mean at either .05 or .02 level.

\*\* Significantly different from Standard Mean at either .01 or .001 level.

The mean MD score for the faked-administration was significantly higher than that for the standard administration ( $p < .001$ ). An MD cutoff score of  $\geq 9$  was used and 3 per cent of the standard administration protocols *versus* 60 per cent of the faked-administration protocols were identified as faked in Group I. With a cutoff score of  $\geq 8$ , 5 per cent of Group I standard administration protocols, *versus* 70 per cent faked-administration are so identified.

The MD scale was not as effective in Group II as in Group I. Only 29 per cent had MD  $\geq 9$ , while 37 per cent had MD  $\geq 8$ . In contrast with Group I, Group II Ss used their real names rather than pseudonyms, and had only



the faked-administration. Whether these procedural differences are responsible for the difference in MD effectiveness cannot be determined from the present data. However, it may be noted that Sorenson and Sheldon (8) have found that such procedural factors can sometimes influence the outcomes of faking studies. In view of the fact that Ss were instructed to try to conceal their faking, correct identification of 37-70 per cent faked-administration protocols at the expense of only 5 per cent standard administration false positives (with  $MD \geq 8$ ) makes the Schanberger scale appear worthy of further study.

In conclusion, the "built-in" protection against distortion described in the manual (4) has been shown to be ineffective and Form A joins Form C in the long list of inventories demonstrated to be quite susceptible to faking.

#### D. SUMMARY

The effects of faking instructions on the 16 PF-Form A were investigated. Differences between standard and faked-administration means were significant for 11 of the 16 factors in a group of 60 undergraduate Ss. Schanberger's Motivational Distortion scale was quite effective in detecting faking in the group of 60 Ss who had both standard and faked-administrations, but did not work as well with 41 Ss who had only a faked-administration.

#### REFERENCES

1. AMERICAN PSYCHOLOGICAL ASSOCIATION. Standards for Educational and Psychological Tests and Manuals. Washington, D.C.: APA, 1966.
2. BRAUN, J. R., & LAFARO, D. Fakability of the Sixteen Personality Factor Questionnaire, Form C. *J. of Psychol.*, 1968, 68, 3-7.
3. CATTELL, R. B., & EBER, H. W. Handbook for the Sixteen Personality Factor Questionnaire. Champaign, Ill.: Inst. for Personal. & Ability Testing, 1957.
4. ———. Manual for Forms A and B, Sixteen Personality Factor Questionnaire. Champaign, Ill.: Inst. for Personal. & Ability Testing, 1962.
5. HANDLER, L., & SHRADER, R. R. Faking on the 16 P-F Test. Paper presented at the annual meeting of the American Psychological Association, New York, New York, 1966.
6. RUNYON, R. P., & HABER, A. Fundamentals of Behavioral Statistics. Reading, Mass.: Addison-Wesley, 1967.
7. SCHANBERGER, W. J. A motivational distortion scale for the Sixteen Personality Factor Questionnaire Form A. Paper presented at the annual meeting of the Western Psychological Association, San Francisco, California, 1967.
8. SORENSON, A. G., & SHELDON, M. S. A further note on the fakability of the MTAI. *J. Appl. Psychol.*, 1958, 42, 74-78.

Department of Psychology  
University of Bridgeport  
Bridgeport, Connecticut 06602

## THE USE OF ANALOGIES AND INCUBATION IN OBTAINING INSIGHTS IN CREATIVE PROBLEM SOLVING\*

*New School for Social Research*

---

ROY DREISTADT

---

### A. INTRODUCTION

The modern study of creative thinking and problem solving began with Helmholtz (12, p. 838) when he took the occasion at a dinner in honor of his 70th birthday to speak about his methods of working on scientific problems. He stated that he often had to wait for "happy thoughts" to help him discover the solution to a difficult problem, and that such a solution never came to him while he was fatigued or working at his desk. A long period of intensive preparatory work on the problem was necessary. Later, when he was rested from the fatigue resulting from this work, the good ideas came to him when he awoke the next morning or while taking a walk. On the basis of his experiences in obtaining insight in this manner, Helmholtz suggested the fatigue and rest hypothesis—that putting the problem temporarily aside—helps a person solve it because it gives him a chance to rest and recover from fatigue that resulted from his previous problem solving efforts.

Poincare (7) described how he often worked intensively on a mathematical problem without success, then put the problem temporarily aside, and after some days or weeks or even months he suddenly got the idea of how to solve the problem. His hypothesis was that unconscious work going on while the problem is put aside eventually causes the sudden insight.

The gestalt psychologists emphasized insight in creative work and K. Buhler used the term "aha experience" (12).

Later, Wallas (10) listed four stages of creative thought: (a) *preparation*, which involves gathering data or information, intensive systematic conscious analysis of the problem, and attempts to develop an hypothesis; (b) *incubation*, which is the stage when the problem is put temporarily aside but is thought about from time to time more or less consciously while the person is engaged in other activities; (c) *illumination*, which is the sudden appearance of the idea in a flash when the person exclaims, "I've got it!" or "Eureka!"; and

---

\* Received in the Editorial Office, Provincetown, Massachusetts, on November 6, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

(*d*) *verification*, which involves the testing of the idea and the final working out of the details of the solution. Sometimes the problem is solved during the preparation phase and the immediate insight then leads to the verification phase. Wallas believed that unconscious work goes on during the incubation phase, while the problem is put aside, that helps the person find the solution.

Still later, Patrick made inquiries among 55 lyric poets (5) and 50 pictorial artists (6) and found that they typically used the four stage creative process. Although some wrote a poem on the spur of the moment or immediately sketched a picture of what they happened to see, 72 per cent of the poets and 76 per cent of the artists reported an incubation stage. She added, however, that the incubated idea or mood recurs from time to time during the incubation period.

Finally, Woodworth and Schlosberg (12) developed the hypothesis that putting the problem temporarily aside in incubation breaks up incorrect sets that are not helping the person solve the problem, thus making it possible for the correct set to emerge. They also believed that recovery from fatigue while the problem is put aside makes it easier to obtain the correct idea.

While these studies were done to try to determine how incubation helps the creative process, Stagner and Karwoski (9) theorized that original discoveries are made by the use of analogies and gave several examples of it. Einstein, in answer to Hadamard's questionnaire (3), stated that he thinks with signs and more or less clear images which he voluntarily reproduces and combines in a vague play aimed to be analogous to certain logical connections he is searching for. The present author (1) presented an analysis of the use of analogies and metaphors by scientists in all the major sciences.

The present study is the first experimental study that seeks to determine the effects of analogies and incubation (an interruption) on creative problem solving. It is also a theoretical analysis, which will be presented in the discussion, that seeks to unify the diverse hypotheses mentioned above and to relate them to recent behavior theory (Hull's theory combined with Guthrie's interference theory) and gestalt theory. The unification of these theoretical systems is also approached thereby.

The purpose of this experiment is to test the following three hypotheses: (*a*) does the perception of visual pictorial analogies aid problem solving; (*b*) does putting the problem temporarily aside (incubation) aid problem solving; and (*c*) does the combined effect of putting the problem aside (incubation) and the presence of visual pictorial analogies aid problem solving?

## B. METHOD

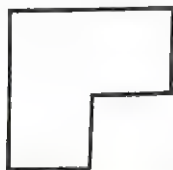
## 1. Subjects

Eighty male and female college students between ages 20 and 35 served as subjects.

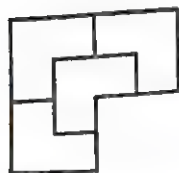
## 2. Procedure

Half of the subjects (40 subjects) were given the farm problem to solve with the instructions to divide the area of the farm into four parts which have the same size and shape. The other half of the subjects (40 subjects) were given the tree planting problem to solve with the instructions to plant 10 trees in five rows with four trees in each row. See Figure 1.

The experiment lasted 20 minutes for each subject. Each subject worked individually on one problem under one experimental condition. Four experimental conditions (with 10 subjects in each condition) were used with each of the two problems.

*The Farm Problem*

DIVIDE THE AREA  
OF THE FARM INTO  
FOUR PARTS WHICH  
HAVE THE SAME  
SIZE AND SHAPE.

*The Tree Planting Problem*

PLANT 10 TREES  
IN 5 ROWS WITH  
4 TREES IN EACH  
ROW.

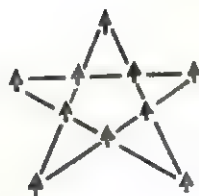


FIGURE 1  
THE FARM PROBLEM AND THE TREE PLANTING PROBLEM  
WITH THEIR INSTRUCTIONS AND SOLUTIONS

In accordance with the four experimental conditions there were (a) control subjects; (b) pictorial analogy subjects; (c) incubation subjects; and (d) combined pictorial analogy and incubation subjects. The control subjects worked on the problem for 20 minutes. The pictorial analogy subjects also worked on the problem for 20 minutes, but there were three pictorial analogies to the problem and its solution, drawn in India ink and water colors on Bristol board, standing in front of them.

The visual pictorial analogies that were shown when the farm problem was used consisted of a green map of the United States with red airplane routes, a yellow-faced clock that is partly covered by a brown radio and whose hands show a quarter to 12 o'clock, and a picture of two graphs drawn on graph paper with green and red lines showing the shape of the problem and one of the pieces in position. The pictorial analogies are shown in Figure 2.

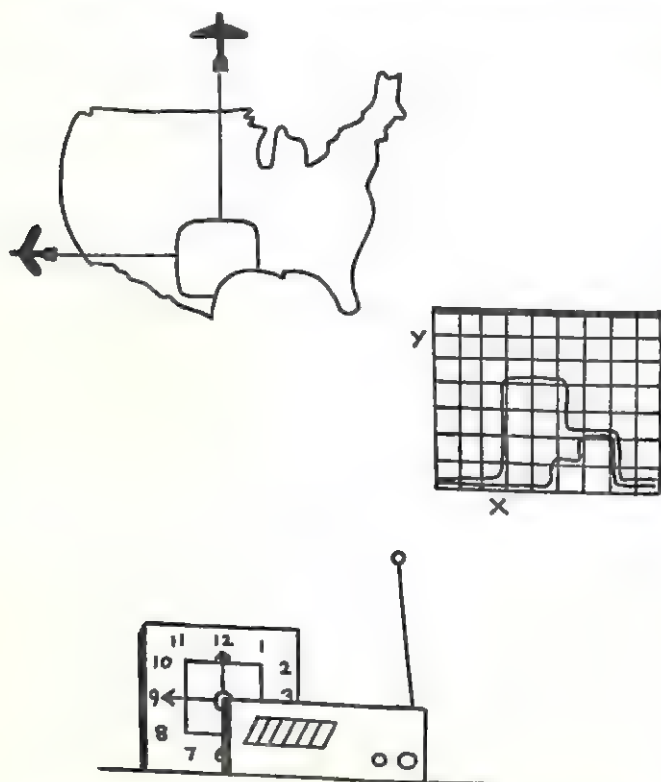


FIGURE 2  
THE VISUAL PICTORIAL ANALOGIES FOR THE FARM PROBLEM



The map of the United States is shaped approximately like the farm problem, and the airplane routes divide the map in similar four parts as would make the solution of the problem. The clock is in the shape of the farm problem (when we disregard the part covered by the radio standing in front of it). If we note the inset design-line of the clock and extend the hands of the clock to touch the rim, and disregard the hands going through the inner area from the center to the inset design-line, then we see the whole solution. The graph picture actually shows the shape of the problem and the other superimposed graph shows one of the four parts of the solution drawn correctly in place.

The visual pictorial analogies that were shown when the tree planting problem was used consisted of a yellow starfish among ordinary-looking blue fish in an aquarium, a Joker card depicting a clown wearing a red star-shaped cap, and a white rocket flying toward the stars (drawn as five-pointed star shapes on a black background). These are shown in Figure 3 and need no explanations, since they are very direct analogies to the problem and its solution. The tree planting problem has a star shaped solution.

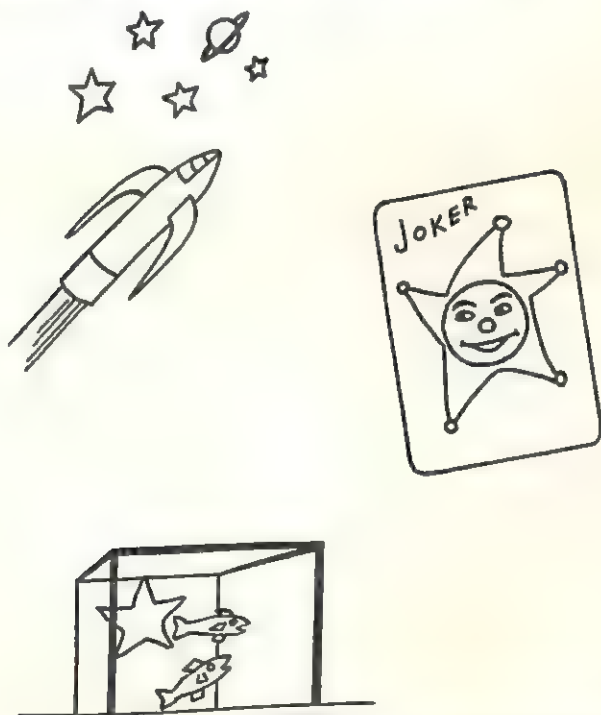


FIGURE 3  
THE VISUAL PICTORIAL ANALOGIES FOR THE TREE PLANTING PROBLEM

The incubation subjects worked on the problem for five minutes (to establish a preparatory stage of problem solving), then they were interrupted for eight minutes (incubation phase) during which they had an easy interpolated task of trying to guess playing cards. After this they were given seven more minutes to try to solve the problem. The following instructions were read to the incubation subjects before they started to work on the problem: "It has been found that when trying to solve a problem, it often helps the person solve the problem if he 'puts the problem aside' for awhile and then continues on it later. You have five minutes in which to solve the problem, then a break of eight minutes, and then seven more minutes to solve the problem."

The combined pictorial analogies and incubation subjects worked on the problem under the same conditions as the incubation subjects except that they had the pictorial analogies (of the problem they were working on) in front of them all the time. They received the same instructions as the incubation subjects.

Since the subjects were all randomly selected, they are likely to have been equated in ability to solve the problem. All the subjects that solved the problem during the first five minutes in any of the four experimental conditions were eliminated because in the two experimental conditions in which there was an incubation phase, the solution would be reached before the incubation. Actually, only one subject (not one of the 80 subjects used) solved the farm problem in less than five minutes. This was one of the pictorial analogy subjects. No subject solved the tree planting problem in less than five minutes.

The farm and tree planting problems were chosen for this experiment because they need no specialized knowledge to solve, because they are tricky puzzle problems, and because one does not often find a person who has already heard of these problems or even others specifically like them. This last point means that these problems are not likely to be solved straightaway by simple transfer of training from previously learned similar tasks. For these reasons, solving the problems of this experiment can be equated to creative thinking and problem solving involved in making original discoveries. The experiment is thus a miniature model of the real thing.

The scores given to solutions and partial solutions for the farm problem were as follows: one point for curved lines or circles, two points for diagonal lines and triangles, three points for rectangles and squares, four points for three correct shapes but placed incorrectly so that the fourth part has the wrong shape, and five points for the correct solution.

The scores given to solutions and partial solutions for the tree planting

problem were as follows: one point for curved rows and separate rows; two points for rows that touch each other (as in making a triangle, etc.); three points for rows crossing each other (as in crosses, letter H's, etc.); four points for rows that touch and cross, but incorrectly; and five points for the correct solution.

The subjects who had the pictorial analogies in front of them (with and without incubation) and solved the problem were asked if they thought that the pictures helped them and, if so, to describe how they helped. This was to determine how conscious and unconscious they were of being helped by the analogies. Introspections were obtained from all subjects after the experiment.

An analysis of variance was done for the four experimental conditions with the farm problem and another analysis of variance was done for the four experimental conditions with the tree planting problem.

### C. RESULTS

The results for the four experimental conditions for the farm problem are shown in Table 1. It can be seen that the scores of the solutions and partial solutions of the subjects who had pictorial analogies and incubation add up to the highest total score which is 45. Seven out of 10 subjects solved the problem completely and one came close to solving it. The second highest total score was 33 and this was for the pictorial analogies subjects. Here three subjects out of 10 solved the problem completely. Most of these subjects said afterwards that they did not take time to look at the pictures carefully even though they noticed that they were there. The total score for the control subjects (who worked on the problem without incubation and without pictorial analogies in front of them) was 28. Only one control subject out of 10 had a

TABLE 1  
THE SOLUTION SCORES OF 40 SUBJECTS FOR THE FARM PROBLEM

No analogies		Analogies	
No incubation	Incubation	No incubation	Incubation
2	2	2	5
3	2	5	5
2	1	2	4
2	3	2	5
2	2	5	5
3	4	3	2
5	2	2	5
3	2	4	5
4	5	3	5
2	1	5	4
Total 28	24	33	45

complete solution. The incubation subjects (without pictorial analogies) did the most poorly and their total score was only 24. One of these subjects solved the problem and another came close to it.

The analysis of variance for the farm problem in Table 2 shows that the effect of the pictorial analogies on the solutions was significant at the 1 per cent level ( $F$  was 12.52). The effect of incubation alone was not significant. Finally, the interaction effect of pictorial analogies and incubation was significant at the 5 per cent level ( $F$  was 4.74).

TABLE 2  
ANALYSIS OF VARIANCE OF THE EFFECT OF ANALOGIES AND INCUBATION ON  
THE SOLUTIONS OF THE FARM PROBLEM OF 40 SUBJECTS

Source of variation	<i>df</i>	Mean square	<i>F</i>
Between groups			
Pictorial analogies	1	16.90	12.52**
Incubation	1	1.60	1.19
Interaction (Pictorial Analogies and Incubation)	1	6.40	4.74*
Within groups	36	1.35	
Total	39		

\* Significant at the 5 per cent level.

\*\* Significant at the 1 per cent level.

When working on the farm problem and trying to divide the area into four parts having the same size and shape, most subjects tended to develop the set or direction of dividing the area of the problem in half and then of further dividing these areas again in half in some way. Then they saw that the sizes and sometimes also the shapes of the parts (depending on how they divided the areas) were not the same. They then tended to persist triangulating the problem in various ways.

The results for the four experimental conditions for the tree planting problem can be seen in Table 3.

The total score from solutions and partial solutions for the subjects who had combined pictorial analogies and incubation was 47. This was the highest total score. Eight out of 10 subjects solved the problem. The pictorial analogies subjects had the second highest total score which was 42. Of these subjects, six out of 10 solved the problem and another two almost solved it. The incubation subjects had a total score of 28 and the control subjects had a slightly lower total score of 26.

The analysis of variance for the tree planting problem (see Table 4) shows that the effect of the pictorial analogies on the solutions was significant above the 1 per cent level ( $F$  was 32.59). This high level of significance is probably

TABLE 3  
THE SOLUTION SCORES OF 40 SUBJECTS FOR THE TREE PLANTING PROBLEM

No analogies		Analogies	
No incubation	Incubation	No incubation	Incubation
3	2	5	5
1	3	5	5
2	2	2	3
2	5	5	5
4	3	2	5
3	2	4	5
2	2	5	5
3	4	4	5
3	2	5	4
3	3	5	5
Total 26	28	42	47

due to the fact that the solution of this problem can readily be perceived as a simple familiar star-shaped whole with the aid of the pictorial analogies. Therefore, one would not expect in the case of the tree planting problem that incubation combined with pictorial analogies could help subjects do very much better. And so we are not surprised when the interaction effect of pictorial analogies and incubation turns out to be small and not significant. Also, the effect of incubation alone was not significant.

TABLE 4  
ANALYSIS OF VARIANCE OF THE EFFECT OF ANALOGIES AND INCUBATION ON  
THE SOLUTIONS OF THE TREE PLANTING PROBLEM OF 40 SUBJECTS

Source of variation	df	Mean square	F
Between groups			
Pictorial analogies	1	30.63	32.59*
Incubation	1	1.23	1.31
Interaction (Pictorial Analogies and Incubation)	1	.22	.23
Within groups	36	.94	
Total	39		

\* Significant at the 1 per cent level.

When working on the tree planting problem subjects usually had difficulty developing any directions at all; they just put the trees anywhere at first hoping to get an idea. Then they would see that, since they had but 10 trees and had to make five rows with four trees in each row, it would be necessary for rows to meet or to cross. In other words, some trees would have to do "double duty."

The verbal reports obtained from the 24 subjects who solved either the farm problem or the tree planting problem and had the pictures of the appro-



priate pictorial analogies in front of them indicate that nine doubted that the pictures helped them and 15 thought that the pictures helped them or may have helped them but they usually were vague as to how the pictures helped. Examples of vague answers were "There was something about the shape of the clock that gave me the idea," "I think seeing the picture of the starfish *might* have helped me," and so on.

#### D. DISCUSSION

The first hypothesis, that visual pictorial analogies help solve problems, was confirmed. Three pictures were in front of a subject as he worked on a problem. These pictures were analogies to the problem he was working on and its solution because they had *similar structures (forms) or groupings to the problem and its solution, while the content or subject-matter of the pictures was different from the problem*. The use of analogies in solving problems involves *transfer with transformation of what is transferred*. The difference in the subject-matter of the analogy is disregarded. The ability and training to disregard the difference in subject-matter of the analogy involves making the subject-matter of the analogy more like *ground* and making the structure of the analogy more *figural*, thereby *abstracting* the structure of the analogy. The part of the abstracted structure of the analogy which is similar to the structure of the problem is matched with it, and is given the subject-matter of the problem; and the rest of the abstracted structure of the analogy is also given the subject-matter of the problem and this is the solution of the problem. In this way the analogy used in solving the problem is *not only transferred to the problem, but the analogy is transformed. There is transfer with transformation of what is transferred*.

The pictorial analogies to each of the two problems used in this experiment contain the structure of the problem and its solution (or at least part of its solution). For example, the farm problem can be solved by using the clock as an analogy because what can be seen of the shape of the clock looks like the problem, while the inner-design with the hands showing a quarter to 12 o'clock looks like the solution. For the tree planting problem, the star shape of any of its analogies depicts—by its shape as a whole—the shape of the solution; and by the points of the angles of the star it shows five rows of such points with four points in each row and a total number of 10 points, which is the statement or specification of the problem. Hence here, too, the problem with its solution is represented in the analogy. The author used these two problems to balance the design of the experiment. The farm problem involves the use of an analytic analogy because we have the whole and use the analogy to discover the parts.

This is like the problem in physics of trying to discover the components of the atom, which was solved roughly at first by using the solar system as an analogy. The tree planting problem involves the use of a synthetic analogy because we have the parts and use the analogy to discover *what* whole they make. This is like the problem in chemistry where John Newlands anticipated the discovery of the Periodic Table (that when the elements are lined up by weight every eighth element has properties that repeat) on the basis of an analogy to music in which every eighth tone repeats. He called his discovery the Law of Octaves in chemistry.

The author has previously (1) analyzed and classified analogies, given examples of analogies that have been used in the sciences, and developed a theory of the use of analogies in science which is part of the theory that will be utilized to explain the results of this research.

The verbal data obtained from the subjects that had the pictorial analogies in front of them and solved the problem indicate that, although far more of these subjects solved the problem than did the subjects who had no pictorial analogies in front of them, they usually were either not conscious that the analogies helped them or they were only vaguely conscious of how the analogies helped them. This is in agreement with the results of an experiment Maier performed (4). Maier found that his subjects solved the pliers and pendulum problem (whose solution involves swinging a string suspended from the ceiling) much more quickly when he gave them a hint by "accidentally" brushing against one of the strings causing it to swing, though they were not conscious of having been given a hint. The present author interprets Maier's "hint" in the following way: when the experimenter brushed against the string and made it swing, he actually did part of the problem for the subject, since the string made to swing by the experimenter was the one needed to be swung to solve the problem. A visual pictorial analogy that would probably help subjects solve the plier and pendulum problem is a picture of a child on a swing or of a grandfather clock which has a swinging pendulum.

The perceptions and thoughts of individual subjects no doubt are different in regard to their conscious imagery. Galton (2) found imagery in most scientists to be poor, but Hadamard (3) found that types of imagery used in thinking vary greatly among individual scientists and mathematicians. We can suppose that there is a relationship between being conscious of using (and being aided by) visual analogies and the tendency to use visual imagery in thought. This is a phenomenological problem that has been partially explored by Hadamard who found many individual differences when he gave mathe-

maticians questionnaires about how they think. For example, Norbert Wiener thought either with or without words, George D. Birkhoff thinks visually, and G. Polya thinks with visual diagrams or "a right word" or even puns.

The physiological correlates of visual imagery in thinking have been studied by Walter (11). Walter found correspondences between individual differences of visual imagery and alpha brain waves. When EEG recordings are taken over the visual area with the eyes closed and the mind at rest, most people show an alpha rhythm. When the person opens his eyes or makes a mental effort, as when doing mental arithmetic, the alpha rhythm disappears and is replaced by a beta rhythm. This is the R type whose alpha rhythms are responsive. Walter also found individuals with persistent alpha rhythms that are hard to block with mental effort, who tend to auditory, kinaesthetic or tactile perceptions rather than visual imagery. In these persons the alpha rhythms continue even when the eyes are open and the mind is active or alert. These persons with persistent alpha activity belong to the P type. He also identified people who have EEG's in which no significant alpha rhythms are found even when the eyes are closed and the mind is resting. This is the M type, for minus; they are persons whose thinking processes are conducted almost entirely in terms of visual imagery. The R type of person can combine data from the various sense organs more readily than can the M or P types. Two-thirds of ordinary normal people selected at random have been found to belong to the R type, and the remaining third are about evenly M and P types. The proportion of M types is usually *higher among science students* than among art students.

Walter's last finding suggests that science students (and scientists), since they are predominantly M types, would tend to use *visual pictorial* types of analogies in their creative thought processes.

The second hypothesis was to determine the effect of putting the problem aside temporarily on problem solving. The incubation subjects were interrupted, and put the problem aside, to mimic operationally the incubation phase (which normally is usually—but not always—a self-imposed interruption). This interruption is equated to an incubation phase because it comes after intensive but unsuccessful attempts to solve the problem (the preparation phase) and because it is only a temporary interruption. The interruption, by interrupting the subject's activities, breaks up his sets. It obviously does not, *by itself*, help the subjects discover better sets. The subjects who had incubation but no pictorial analogies in front of them had little opportunity, during the short interruption and in the restricted visual environment without the pic-

tures, to find the right set or direction for solving the problem. These subjects tended to return to their previous sets. They would often say, "I know I've tried this before, but I cannot think of anything else to do." Because of the interruption, these subjects also worked on the problem for a shorter length of time. These may be the reasons why incubation by itself had no significant effect on solving either the farm problem or the tree planting problem. The interruption of the incubation phase may help because it breaks up sets or directions and because it gives rest from fatigue, but it is also apt to work against a person because he spends less time working on the problem. These results indicate that the incubation (interruption) phase in creative problem solving serves as something more than merely to break up sets and give rest to recover from fatigue. The person also finds the right set or direction during incubation and this gives the illumination.

The third hypothesis tested the combined effect of pictorial analogies and incubation on problem solving. This resulted in the greatest number of solutions (and good partial solutions) of the two problems, and indicates that incubation enhances the effect of the pictorial analogies. This interaction effect was significant for the subjects with the farm problem, but was not significant for the subjects with the tree planting problem. This latter finding probably occurred because the pictorial analogies to the tree planting problem were direct enough to be perceived almost as well when a subject was constantly working on the problem as when he put the problem aside temporarily with the pictures still in front of him. Here incubation is less likely to be needed in order to enable a person to give sufficient attention to the pictures so that he can be aided by them. In real life situations visual stimulus-patterns that can be used as analogies to solve a problem are not usually to be found so close at hand and so incubation is generally helpful because it brings a person back into contact with the larger, more enriched behavioral environment with its manifold and varied patterns of stimulation where he may find a pattern of stimulation—which was not in the immediate environment while he worked on the problem—that he can use as an analogy for solving the problem.

The author proposes that incubation—with the breaking up of sets and the rest from fatigue—should be interpreted in the combined language of gestalt and behavior theory (combining Hull's inhibition and Guthrie's interference theories).

Reminiscence has been explained by the development of  $I_R$  during learning (especially with mass practice) which causes some forgetting, but then dissipates so that a person recalls more later on than immediately after learning.



This effect is due only to  $I_R$  and dissipation of  $I_R$ . Incubation, in creative problem solving, is more complex. We must bear in mind that when a person discovers something new he is not learning what is shown to him; he is learning something new *as he discovers it*. A child today can learn about the law of gravity from his teacher. Newton learned about the law of gravity by discovering it. Yet Newton used what he recalled from his past learning to discover it.

We follow Hull's formula, that habit strength ( $sH_R$ ) multiplied by drive ( $D$ ) equals reaction potential ( $sE_R$ ). From this we can assert that the habit strength of a response (determined by frequency) multiplied by drive intensity causes the response (or response tendency—as when one puts oneself into a posture or stance in readiness for making an intended or desired response) to have a "set," by making the response (or response tendency) rigid or "set" in its motion (or motion tendency), respectively. Hence, any response or response tendency has a "set" or "direction."

Simple problems—not used in this experiment—can be solved with very little trial and error because there are only a very limited number of responses possible to the problem, and one of these few possible responses solves the problem. Complex problems—as the two used in this experiment—have a great many possible responses if the person has to try them all by trial and error to see which response solves the problem. The person may find that one of his predominant responses that gets transferred to the problem solves it. When this happens he solves the problem rapidly. But if the person has no predominant response from previous learning that gets transferred to the problem and solves it, it may take him a very long time to solve a complex problem. This is the problem-situation in this experiment.

Spence (8) theorized that strong motivation is good for learning simple tasks rapidly, but with difficult or complex tasks strong motivation at first tends to interfere because it energizes wrong responses that are predominant, but subsequently (as the wrong responses are inhibited) strong motivation helps learn the task. The subjects with strong motivation working on complex tasks surpass the subjects with weak motivation working on complex tasks. It should be pointed out that a simple task may be difficult (if it involves much physical exertion), but a task may be difficult because it is complex with many parts to learn or memorize. A task may be difficult for both reasons. The present author defines *problems as incompletely given tasks*. Then simple problems (which are only slightly incomplete tasks) are solved rapidly when motivation is strong. This author further hypothesizes that *with complex problems strong motivation is needed to cause the person to work intensely on the prob-*



*lem—preparation phase—and then to cause fatigue and induce incubation but yet to cause the person (in a less intensive manner) to keep coming back to the problem, which makes him likely to solve it eventually.*

In the preparation phase, when working on one of the problems in this experiment, the subjects first made responses which were predominant in relation to the stimulus-patterns of parts of the problem-situation, and they therefore sometimes succeeded in solving parts of the problem. But usually the subjects, in working on parts of the problem, *worked on the kind of parts that already prevented them from solving the problem.* For example, many subjects first divided the area of the farm problem in half (with a diagonal line) and then tried unsuccessfully to divide these two equal parts into halves again to obtain four parts of equal size and shape. They no doubt transferred to the problem the response of bisecting the area because they saw it as being like the geometrical problems that they had previously studied in geometry class which are frequently solved by bisecting areas. Typically, the subjects tried their predominant responses again but with some modifications or by combining the responses in various ways. For instance, subjects who divided the area in half often tried to halve these two parts in various ways. Some subjects even erroneously modified their straight lines into curved lines. The subjects often returned to responses which they had tried already.

Thus, when unsuccessfully working on the problem with strong motivation, in the preparation phase, a person's predominant responses—since they tend to be repeated with modifications—eventually undergo strong and massive fatigue which produces reactive inhibition ( $I_R$ ) of all the predominant responses to the problem. This reactive inhibition causes him to slow down and temporarily stop working on the problem, which is to say that it brings about or *induces the incubation phase.* In the incubation phase the person may stay in the presence of the problem (as in this experiment) or walk away, but in any case he carries the problem around in his head, and sees himself not making his previous (predominant) responses to the problem which therefore undergo some conditioned inhibition ( $sI_R$ ). He usually also makes other responses, as he behaves in other situations during incubation. Some of these responses interfere with the previous responses that have undergone  $I_R$ , and produce an *active* conditioned inhibition of them (by doing something else) just as Hull's *active* conditioned inhibition of them (holding back or withdrawing implicitly) produces nonresponding response (*passive* conditioned inhibition ( $sI_R$ )). Both are really conditioned inhibition ( $sI_R$ ). The  $I_R$  plus  $sI_R$  is the total inhibition which extinguishes the previous responses. Later, *as incubation continues*, he gradually recovers from fatigue;  $I_R$  dissipates. This is why the incubation phase, as it goes on, makes him feel

rested from fatigue or refreshed as Helmholtz stated. The conditioned inhibition of the previous (predominant) responses remains, making these responses weaker (attenuated) when they spontaneously recover (as reactive inhibition dissipates). The sets of the previously predominant responses are broken up because they are weakened by having undergone some conditioned inhibition. This is the Woodworth and Schlosberg "breaking up of sets" effect in incubation.

Insofar as some of the conditioned inhibition is produced by interference from the responses he is making to various situations during incubation, these situations (stimulus-patterns) also determine *which new responses he makes to the problem*. The same responses he makes to the situations he behaves in during incubation are *also* made to the problem, and so new sets or directions for solving the problem develop. These responses, since they belong to the situations he behaves in during incubation, are made "aside," haphazardly, and unconsciously to the problem (as Poincare stated). If he behaves in situations whose stimulus-patterns *or* whose stimulus-patterns and his responses to them can be used as *analogies* to solve the problem, he solves the problem suddenly by insight or illumination.

The use of analogy involves either (a) matching the stimuli of the problem to *some of the stimuli* of one of the stimulus-situations experienced during incubation and then making responses to the problem that imitate (copy) the rest of the stimulus-pattern of that stimulus-situation (as when Newlands saw that some elements repeat like the tones in music and then lined them all up into octaves like the piano keyboard), or (b) matching the stimuli of the problem to the stimuli of one of the stimulus-situations experienced during incubation and then making the same responses to the stimuli of the problem as he makes to the stimuli of this stimulus-situation (as when Archimedes saw the water overflow when he stepped into the bath and then put the crown into a tray of water to measure the overflow).

### E. SUMMARY

This was an experimental study of the use of analogies and incubation in creative problem solving. Subjects were given the farm problem or the tree planting problem to solve. For each problem there were control subjects who just worked on the problem, incubation subjects who put the problem aside temporarily and later continued to work on it, pictorial analogy subjects who had pictorial analogies of the problem in front of them while they worked on the problem, and combined pictorial analogy and incubation subjects who had the pictorial analogies in front of them and also had an incubation phase in

which they put the problem aside temporarily. It was found that pictorial analogies significantly aided the subjects in solving the two problems and that the interaction effect of pictorial analogies and incubation significantly aided the subjects in solving the farm problem (which had a more complex solution). The results were discussed and explained from a unifying viewpoint involving several theories.

## REFERENCES

1. DREISTADT, R. An analysis of the use of analogies and metaphors in science. *J. of Psychol.*, 1968, 68, 97-116.
2. GALTON, F. *Inquiries into the Human Faculty and Its Development*. London, England: Macmillan, 1883.
3. HADAMARD, J. *The Psychology of Invention in the Mathematical Field*. Princeton, New Jersey: Princeton Univ. Press, 1949.
4. MAIER, N. R. F. Reasoning in humans: II. The solution of a problem and its appearance in consciousness. *J. Comp. Psychol.*, 1931, 12, 181-194.
5. PATRICK, C. Creative thought in poets. *Arch. Psychol.*, 1935, 26, No. 178.
6. ———. Creative thought in artists. *J. of Psychol.*, 1937, 4, 35-73.
7. POINCARÉ, H. Mathematical creation. In B. Gheselin (Ed.), *The Creative Process*. Berkeley, Calif.: Univ. Calif. Press, 1952. Pp. 33-42.
8. SPENCE, K. W. A theory of emotionally based drive (D) and its relation to performance in simple learning situations. *Amer. Psychologist*, 1958, 13, 131-141.
9. STAGNER, R., & KARWOSKI, T. F. *Psychology*. New York: McGraw-Hill, 1952. P. 419.
10. WALLAS, G. *The Art of Thought*. New York: Harcourt Brace, 1926.
11. WALTER, W. G. *The Living Brain*. New York: Norton, 1953.
12. WOODWORTH, R. S., & SCHLOSBERG, H. *Experimental Psychology*. New York: Henry Holt, 1954. Chapter 26.

256 East 89 Street  
New York, New York 10028



## THE EFFECTS OF PUNISHMENT ON FOOD PELLET DISCRIMINATION\*<sup>1</sup>

*Department of Psychology, The University of South Alabama*

---

WILLIAM F. VITULLI

---

### A. INTRODUCTION

The discriminative properties of food pellets have been of continuing interest with respect to their effects in schedules of reinforcement since the observation made by Skinner (11). Discriminative stimulus control by the type of food pellet was demonstrated by Cruse, Vitulli, and Dertke (4). Noyes sucrose and standard pellets controlled rates of responding in a multiple fixed-ratio 30 fixed-interval three-minute schedule of reinforcement. Also, discriminative stimulus induction across a food pellet continuum of different sucrose concentrations was found by Vitulli (13). Eleven mixtures of Noyes sucrose and standard food pellets were used as discriminative stimuli under a two-ply multiple schedule. Response rates following each pellet concentration were controlled by the presence of the larger proportion of sucrose or standard food. Loss of discriminative control occurred when pellets composed of .50 sucrose and .50 standard food preceded the schedule components of the multiple. In each of these studies the stimulus properties of the primary reinforcers were sufficient to produce operant discrimination.

Acquired discriminative properties of aversive events, such as shock or intense noise, have been studied, for example, by Azrin (1) and by Holz and Azrin (7). Their paradigm included associating the aversive stimulus with positive primary reinforcement. The focus of analysis was the subsequent stimulus control exerted by the aversive stimulus.

The purpose of the present investigation was to assess the effects of pairing an aversive stimulus with distinctive primary reinforcers. Subsequent discriminative properties of the primary reinforcers would be measured. It was reasoned that by consistently associating shock with one type of pellet and not with another, in a fixed-ratio (FR) schedule of reinforcement having the same

---

\* Received in the Editorial Office, Provincetown, Massachusetts, on November 8, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> The equipment used in this research was funded in part by Grants 01-6379 and 01-6391 from the Research Committee of the University of South Alabama. Appreciation is expressed to the following students from the Work Study Program for their invaluable assistance: Robert Battles, Kathryn Buntyn, and Daniel Koch.



response requirements (30) for both pellets, postreinforcement rate differences would occur between pellets. Rate differences would occur not only when shock was present but also when shock was removed. If the sucrose and standard pellets were presented in a random sequence for a given session, the rat could not learn which pellet was "coming next." The animal could, however, learn to discriminate between the shock-paired *vs.* the nonshock-paired pellet types. This procedure separated the discriminative properties from the reinforcing properties of the pellet. Skinner (11), Keller and Schoenfeld (8), and Terrace (12) discuss at length the theoretical and empirical differences between these properties. The contingencies thus described may be referred to as a *mult* FR 30 FR 30 schedule of reinforcement. The accompanying stimuli are the pellet types and the presence or absence of shock.

## B. METHOD

### 1. Subjects

Two male Holtzman albino rats (S3 and S4) with no previous training were used. The animals were approximately three months old and weighed between 260 and 275 grams under conditions of free food and water at the beginning of the experiment. They were reduced to 85 per cent of their free-feeding weights and were maintained at from 75 per cent to 85 per cent of free-feeding weight throughout the experiment.

### 2. Apparatus

A Grason-Stadler one-lever rat chamber was used. The chamber received pellets from two Davis pellet dispensers. One dispenser presented Noyes stock 45 mg standard food pellets. The other dispenser presented Noyes stock 45 mg sucrose food pellets. The dispensers were mounted on a wall directly above and outside of the experimental enclosure. A Y-shaped surgical tubing arrangement extended from the dispensers to the food hopper in the chamber. Refer to Vitulli (13) for further description of this apparatus. A third dispenser (Grason-Stadler) was located inside the experimental enclosure and contained no pellets. The Grason dispenser operated whenever either of the Davis dispensers operated to mask their possible differential sounds. A Foringer white noise generator was connected in series with the chamber speaker. Its continual operation masked possible extraneous sounds occurring outside of the experimental enclosure. A Grason-Stadler shock generator was the source of scrambled electric shock stimuli. Conventional programming and switching equipment was used.

### 3. Procedure

After the animals were reduced to 85 per cent of their free-feeding weight, S3 was magazine trained with standard pellets and S4 was magazine trained with sucrose pellets. Both organisms were then rewarded for every lever response by the same type of pellet which was used during magazine training. The response ratio was gradually increased for both rats to an FR 30 over the following 12 sessions. After five sessions on FR 30 with standard pellets for S3 and sucrose pellets for S4, each subject received both types of pellets under the *mult* FR 30 FR 30 schedule. At this point the accompanying stimuli for FR 30 components in the multiple schedule were different type food pellets. After 10 sessions, a mean shock intensity of .55 milliamps (ma) and 100 milliseconds (msec) was paired simultaneously with the occasion of each standard pellet in the *mult* FR 30 FR 30 for S3. The onset of the shock occurred at the same time as the dropping of the food pellet in the hopper. A mean shock intensity of .54 ma and 80 msec was paired simultaneously with the occasion of each sucrose pellet in the *mult* FR 30 FR 30 for S4. Again, the onset of the shock occurred at the same time as the dropping of the food pellet in the hopper. During blocks of sessions in which shock was presented, each subject received both types of pellets. One type of pellet was consistently associated with shock and the other pellet was never associated with shock.

Daily sessions lasted for an average of 60 pellets or one hour, whichever came first, for each subject. The animals were not run on the weekends (Saturday and Sunday). The pellet sequence under the multiple schedule was random and proportionally stratified for each session so that both animals received an average of 50 per cent sucrose and 50 per cent standard pellets per session. Refer to McNemar (9) for further description of proportional stratification. The sequence was programmed by aperiodic input to a Scientific Prototype probability generator set for .50 output. Blocks of sessions alternated with respect to the presence or absence of shock correlated with the respective pellet type per subject. The total program for mean shock intensity and duration per pellet for S3 and S4 is presented in Table 1. The sequence of sessions and shock properties per block refers to the successive order in which they occurred. Variations in mean shock intensity and duration were presented in early sessions to determine those values at which the rates of each subject would stabilize. At Block 6 it was found that both animals were emitting stable rates at an absolute shock intensity of .40 ma and 50 msec duration. These values were then used without exception for the remaining shock sessions for both animals.

TABLE 1  
 PELLET TYPE, MEAN SHOCK INTENSITY AND DURATION PER BLOCK OF SESSIONS IN THE  
 MULTIPLE FIXED-RATIO 30 FIXED-RATIO 30 SCHEDULE FOR S3 AND S4

Blocks of sessions	No. of sessions	Shock intensity and duration with pellet type	
		S3: Standard	S4: Sucrose
1	10	No Shock	No Shock
2	33	.55ma-100msec	.54ma-80msec
3	10	No Shock	No Shock
4	10	.33ma-60msec	.30ma-50msec
5	5	No Shock	No Shock
6	5	.40ma-50msec	.40ma-50msec
7	5	No Shock	No Shock
8	5	.40ma-50msec	.40ma-50msec
9	5	No Shock	No Shock
10	5	.40ma-50msec	.40ma-50msec
11	5	No Shock	No Shock
12	5	.40ma-50msec	.40ma-50msec
13	5	No Shock	No Shock
14	5	.40ma-50msec	.40ma-50msec

### C. RESULTS

Figure 1 presents stable absolute rates for the terminal three sessions partitioned from each block for S3 and S4. The rates indicate that the pairing of shock with standard pellets for S3 and sucrose pellets for S4 lowered postreinforcement rates consistently in comparison with rates following pellets not associated with shock. Blocks of sessions in which shock was removed (Blocks 3, 5, 7, 9, 11, and 13) indicate that postreinforcement rates were consistently higher following the pellets previously associated with shock in comparison with rates following the pellets never associated with shock. The rates following pellets never associated with shock (sucrose for S3 and standard for S4) tended to be lower in those sessions during which shock was associated with the alternate pellet type (Blocks 2, 4, 6, 8, 10, 12, and 14). The rates were consistently higher in those sessions during which no shock was associated with the alternate pellet type, as indicated in Figure 1 (Blocks 1, 3, 5, 7, 9, 11, and 13).

### D. DISCUSSION

The association of an aversive stimulus with positive primary reinforcement in the present study tended to alter the discriminative properties of the primary reinforcing stimuli. Comparisons between stable response rates following the pellets previously paired with shock with stable response rates following the pellets never paired with shock, within the same organism, indicate that the former pellets controlled consistently higher rates than the latter pellets.

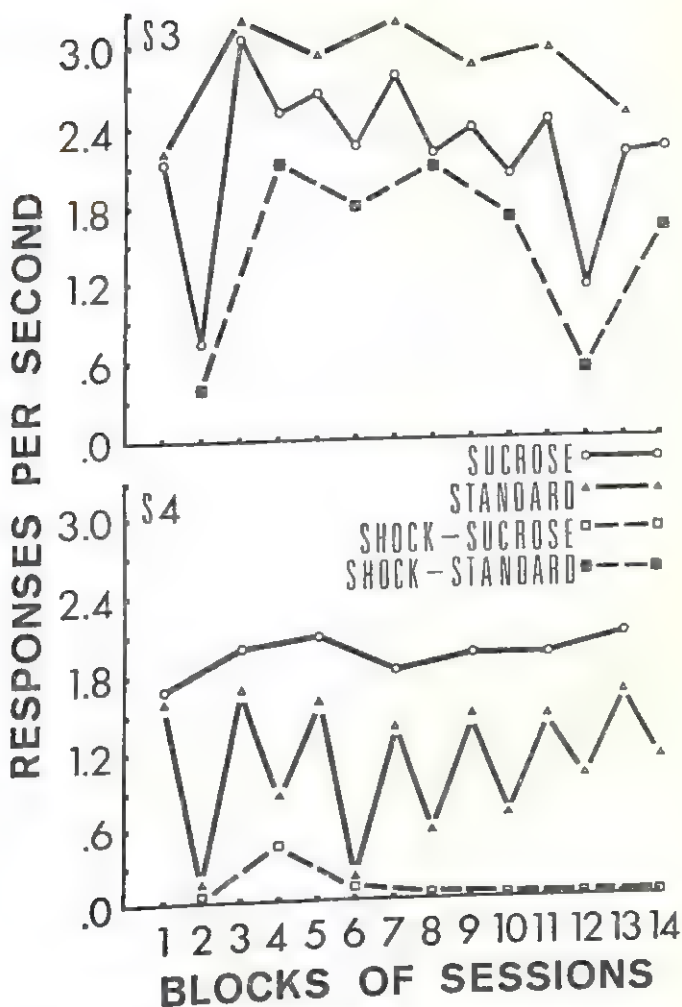


FIGURE 1  
STABLE ABSOLUTE RESPONSE RATES FOLLOWING SUCROSE AND STANDARD PELLETS IN THE  
*mult* FR 30 FR 30 SCHEDULE FOR THE TERMINAL THREE SESSIONS OF EACH  
SUCCESSIVE BLOCK FOR SUBJECT 3 AND SUBJECT 4

There was no evidence that a pellet became a conditioned aversive stimulus because of its previous association with shock under the conditions of the present investigation. Conversely, postshock session rates were facilitated rather than suppressed. These results are consistent with the findings of Garcia, McGowan, Ervin, and Koelling (6). They report that when the consequence of

eating is immediate peripheral pain, produced by a shock of .2-sec and 2.0 ma, rats do not display a reduced preference for either gustatory or nongustatory attributes of the food. Garcia *et al.* (6) used a respondent conditioning paradigm. Possible changes in performance were measured by the latency between the presentation of food and the consummatory response. The flavor of the pellet did not acquire negative CS properties.

Behavioral contrast describes the postshock session rate relationships produced between the pellets in the present investigation. Reynolds defined contrast as "when the change in the rate of responding during the presentation of one stimulus is in a direction *away from* the rate of responding generated during the presentation of the other stimulus" (10, p. 57). Postshock rate divergencies between pellets are seen in Figure 1. The schedule of reinforcement (*mult* FR 30 FR 30) appeared to have gained considerable control producing overall high rates during nonshock sessions. A multiple schedule in which components typically produce low rates—e.g., variable-interval components—may reveal greater contrast under the conditions of the present study.

The presence of response-contingent shock associated with a distinctive pellet had punishing effects. The shock suppressed responding. Azrin and Holz define punishment as a "consequence of behavior that reduces the future probability of that behavior" (3, p. 381). The data in Figure 1 indicate that rates were lowest following pellets paired with punishment. The notion, expressed by Estes (5) and Azrin (2), that punishment suppresses responding but does not permanently weaken it is evidenced in alternate blocks of sessions during which shock was not present. Rates of responding in nonshock sessions tended to rise in direct proportion to the degree of previous suppression.

Major conclusions suggested by the present investigation were that (a) the pairing of a punishing stimulus with primary reinforcement suppressed postreinforcement rates, (b) removal of the associated punishment modified the discriminative properties of the pellet so that postreinforcement rates were relatively higher, and (c) there was no evidence in this study that associating shock with food rendered the food a conditioned aversive stimulus. A systematic replication with additional animals, and parametric variations of the food-shock interval, ought to increase the generality and reliability of these conclusions.

### E. SUMMARY

Two rats were trained on a multiple fixed-ratio 30 fixed-ratio 30 schedule of reinforcement. The reinforcers were Noyes standard and sucrose food pellets. The presentation of the two types of pellets was random. The onset of a



shock of approximately .40-ma and 50-msec was paired, simultaneously and consistently, with one type of pellet. Blocks of sessions alternated regarding the presence or absence of the response-contingent shock for each rat. One animal received shock with standard pellets and the other animal received shock with sucrose pellets in the multiple schedule. The presence of shock lowered postreinforcement rates of responding following the shock-paired pellet. Response rate decrements were less following the pellet type not paired with shock during the same session. Removal of shock for the pellet previously paired with shock resulted in consistently higher postreinforcement rates as compared with rates following the pellet never paired with shock. The data were interpreted in terms of punishment, behavioral suppression, and contrast.

## REFERENCES

1. AZRIN, N. H. Some effects of noise on human behavior. *J. Exper. Anal. Behav.*, 1958, **1**, 183-200.
2. ———. Sequential effects of punishment. In T. Verhave (Ed.), *The Experimental Analysis of Behavior: Selected Readings*. New York: Appleton-Century-Crofts, 1966. Pp. 186-189.
3. AZRIN, N. H., & HOLZ, W. C. Punishment. In W. K. Honig (Ed.), *Operant Behavior: Areas of Research and Application*. New York: Appleton-Century-Crofts, 1966. Pp. 380-447.
4. CRUSE, D. B., VITULLI, W. F., & DERTKE, M. Discriminative and reinforcing properties of two types of food pellets. *J. Exper. Anal. Behav.*, 1966, **9**, 293-303.
5. ESTES, W. K. An experimental study of punishment. *Psychol. Monog.*, 1944, **57** (No. 3), Whole Number 263.
6. GARCIA, J., MCGOWAN, B. K., ERVIN, F. R., & KOELLING, R. A. Cues: their relative effectiveness as a function of the reinforcer. *Science*, 1968, **160**, 794-795.
7. HOLZ, W. C., & AZRIN, N. H. Discriminative properties of punishment. *J. Exper. Anal. Behav.*, 1961, **4**, 225-232.
8. KELLER, F. S., & SCHOENFELD, W. N. *Principles of Psychology*. New York: Appleton-Century-Crofts, 1950.
9. MCNEMAR, Q. *Psychological Statistics* (3d ed.). New York: Wiley, 1963. P. 383.
10. REYNOLDS, G. S. Behavioral contrast. *J. Exper. Anal. Behav.*, 1961, **4**, 57-71.
11. SKINNER, B. F. *The Behavior of Organisms: An Experimental Analysis*. New York: Appleton-Century-Crofts, 1938.
12. TERRACE, H. S. Stimulus control. In W. K. Honig (Ed.), *Operant Behavior: Areas of Research and Application*. New York: Appleton-Century-Crofts, 1966. Pp. 271-344.
13. VITULLI, W. F. Stimulus properties of varied food concentrations. *Psychon. Sci.*, 1967, **9**, 495-496.

Department of Psychology  
The University of South Alabama  
Mobile, Alabama 36608



## VISUAL ACUITY DECREMENT FOR A SUBFUSIONAL INTERMITTENCY WITH TWO DIRECTIONS OF BORDER CONTRAST\*<sup>1</sup>

*Laboratory for the Study of Vision and Related Sensory Processes,  
Michigan State University*

---

S. HOWARD BARTLEY AND RICHARD J. BALL

---

### A. INTRODUCTION

Unique changes in perceived brightness, hue, and saturation are produced by subfusional intermittency rates in the range of about five to 10 cycles per second. The occurrence of brightness enhancement, temporally induced desaturation, and alterations in perceived hue have been extensively studied in this laboratory (1, 2, 3, 4, 6, 7, 9). Temporal manipulation of photic input also has marked effects on a number of other aspects of visual perception (5, 14, 15). One of these is visual acuity.

Bourassa and Bartley (11) found that visual resolution of a gap between two halves of an illuminated target was adversely affected under the same intermittency conditions that produced brightness enhancement. Using a tumbling E target Bartley, Nelson, and Soules (10) found visual acuity decreasing as temporally induced brightness increased. Senders (16, 17, 18), Nachmias (13), and Gerathewohl and Taylor (12) have also shown the unique effects that intermittent stimulation produces on visual acuity.

A recent study by Bartley and Ball (8) more extensively investigated these relationships. Utilizing a Landolt C target, they found maximum acuity decrement at rates around five cycles per second with less decrement at faster and slower rates. A pulse-to-cycle fraction of  $\frac{1}{4}$  produced more decrement than those of  $\frac{1}{2}$  and  $\frac{3}{4}$ , but it was shown that the pulse length of the intermittent stimulation was not the prime factor. Total flux was also shown not to be a prime factor except at very low levels of luminance. In a reverse contrast portion of the experiment, it was found that a black Landolt C on a white background produced much more acuity decrement than a white Landolt C on a black background.

---

\* Received in the Editorial Office, Provincetown, Massachusetts, on November 18, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> Investigation supported in whole by Public Health Service Research Grant NB 05260-05 from the National Institute of Neurological Diseases and Blindness.

## B. APPARATUS

The two targets were Landolt C's which could be randomly presented with the gap-oriented "up," "down," "right," and "left." The black Landolt C was made with India ink on matte white paper, while the white Landolt C was white ink on matte black paper. Both letters presented a high contrast target.

The collimated output of a 750-watt projection bulb controlled by a Variac was used to illuminate this target field. A variable-speed drive Graham-transmission with an episcotister disk provided the intermittency. The episcotister chopped the light beam at a focal point in order to minimize taper.

## C. PROCEDURE

This study is a direct outgrowth of the study (8) previously discussed. Maximum acuity decrement was obtained in that study with a rate of five cycles per second, PCF of  $\frac{1}{4}$ , and luminance of 96 c/ft<sup>2</sup>. Therefore, these conditions of rate, PCF, and luminance were selected for this present study. The purpose of this study was to investigate the temporal induction of acuity decrement as visual angle was systematically decreased for conditions of both black Landolt C on white background and white Landolt C on black background.

The visual angle was changed by varying the distance between the observer and the target. Both Landolt C targets utilized, subtended five minutes of visual angle at 15 feet. These targets were viewed at distances between 7.5 feet and 30 feet so the target visual angle varied from 10 minutes down to 2.5 minutes.

A random order presentation series of 10 Landolt C gap orientations was utilized as a single run for each experimental condition. A forced-choice procedure was used in which the observer had to report the gap as "up," "down," "right," or "left." The graphs presented are averages of 10-trial runs for each condition for each of the two observers.

## D. RESULTS

The results are summarized in Figure 1. The abscissa is the visual angle of the target in minutes of arc. The ordinate is the average number of errors made in a group of 10 randomized Landolt C presentations. The alternate dash-dot line is the Steady I condition which was a luminance of 24 c/ft<sup>2</sup>. The broken line is the Steady II condition with a luminance of 96 c/ft<sup>2</sup>. The solid line is the intermittent condition of rate five cycles per second and

pulse-to-cycle fraction of  $\frac{1}{4}$ . Thus, the Steady I condition is equated in total flux per cycle to the intermittent condition, and the Steady II is equated to the flux density of the pulse in the intermittent condition.

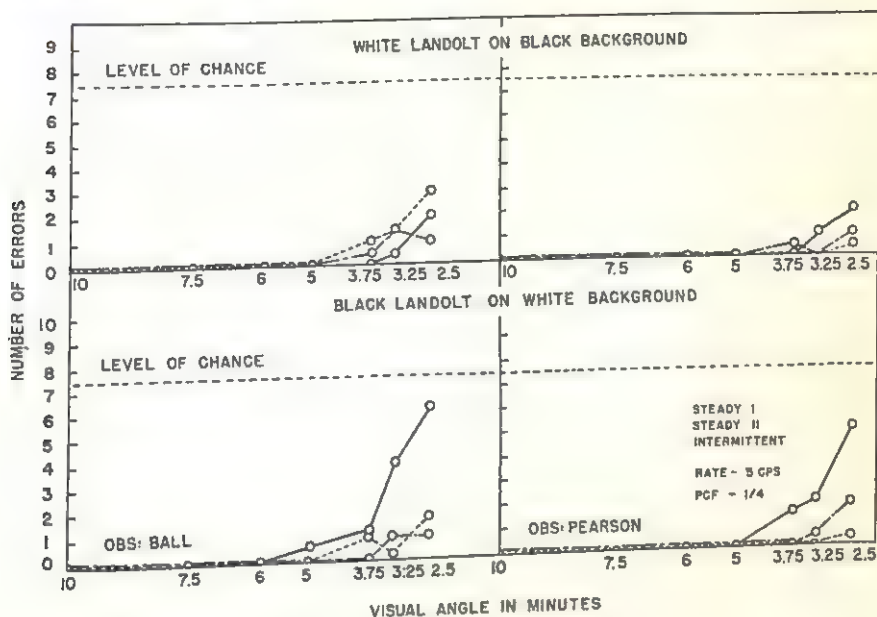


FIGURE 1  
THE AVERAGE NUMBER OF ERRORS IN A 10-TRIAL RUN FOR DETECTION OF THE POSITION OF A LANDOLT C PLOTTED AGAINST THE VISUAL ANGLE OF THE LETTER

Results for both intermittent stimulation (rate 5 cps; PCF of  $\frac{1}{4}$ ) and two levels of steady stimulation are shown. Data for two observers and both black and white Landolt C conditions are presented.

Figure 1 shows that there is very little difference between the two steady conditions, or, that a difference in luminance between 24 c/ft<sup>2</sup> and 96 c/ft<sup>2</sup> is not a prime factor in acuity decrement. This is the case for both the white letter on black background and the black letter on white background. Very few errors are made under any of the steady stimulation conditions and no errors with a visual angle greater than 3.75 minutes.

The intermittent condition does not produce any errors with a visual angle greater than five minutes. However, with visual angles less than five minutes, the acuity decrement rapidly increases with decreasing visual angle for the black Landolt C on white background and is approaching the level of change at a visual angle of 2.5 minutes. For the white Landolt C on black back-



ground, there is very little difference between the intermittency condition and the two steady conditions.

### E. DISCUSSION

This study once again demonstrates the unique effects of intermittency of stimulus input on a perceptual end result—in this case, visual acuity. Previous studies (8, 10, 11) have shown that decrements in visual acuity were produced by intermittency rates in the same range that produce brightness enhancement. In the present study, an intermittency condition previously found to produce visual acuity decrement was utilized to investigate the increase in acuity decrement as visual angle was systematically decreased. It was found that for visual angles of less than five minutes, the number of errors rapidly increased for black targets on a white background. This visual acuity decrement did not occur for white targets on a black background.

It appears plausible that this could be explained on the basis that when a large proportion of the target field is actively stimulating the visual system (black letter on bright white field), the temporally induced effects produced by the intermittent stimulation disrupt the resultant contour processes necessary for visual acuity discrimination. With the reverse situation of only a small proportion of the total target field actively stimulating the visual system (white letter on black background), the temporally induced effects associated with alternating of response mechanisms would not be sufficient to disrupt seriously the resultant perception of borders.

### REFERENCES

1. BALL, R. J. An investigation of chromatic brightness enhancement tendencies. *Amer. J. Optom.*, 1964, **41**, 333-361.
2. BALL, R. J., & BARTLEY, S. H. Effects of temporal manipulation of photic stimulation on perceived brightness, hue and saturation. *Amer. J. Optom.*, 1965, **42**, 573-581.
3. ———. Changes in brightness index, saturation, and hue produced by luminance-wavelength-temporal interactions. *J. Opt. Soc. Amer.*, 1966, **56**, 695-699.
4. ———. Further investigations of color perception under temporal manipulation of photic stimulation. *Amer. J. Optom.*, 1966, **43**, 419-430.
5. ———. The induction and reduction of color deficiency by manipulation of temporal aspects of photic input. *Amer. J. Optom.*, 1967, **44**, 411-418.
6. ———. Brightness and saturation changes resultant from temporal manipulation of spectral photic input for normal and color deficient observers. *J. of Psychol.*, 1968, **68**, 55-61.
7. BARTLEY, S. H., & BALL, R. J. The study of target size in production of brightness, hue and saturation for intermittent photic stimulation in the red end of the visible spectrum. *J. of Psychol.*, 1966, **64**, 241-247.
8. ———. Effects of intermittent illumination on visual acuity. *Amer. J. Optom.*, 1968, **45**, 458-464.

9. BARTLEY, S. H., & NELSON, T. M. Certain chromatic and brightness changes associated with rate of intermittency of photo stimulation. *J. of Psychol.*, 1960, **50**, 323-332.
10. BARTLEY, S. H., NELSON, T. M., & SOULES, E. M. Visual acuity under conditions of intermittent illumination productive of paradoxical brightness. *J. of Psychol.*, 1963, **55**, 153-163.
11. BOURASSA, C. M., & BARTLEY, S. H. Some observations on the manipulation of visual acuity by varying the rate of intermittent stimulation. *J. of Psychol.*, 1965, **59**, 319-328.
12. GERATHEWOHL, S. J., & TAYLOR, W. Effect of intermittent light on the readability of printed matter under conditions of decreasing contrast. *J. Exper. Psychol.*, 1953, **46**, 276-282.
13. NACHMIAS, J. Brightness and visual acuity with intermittent illumination. *J. Opt. Soc. Amer.*, 1958, **48**, 726-730.
14. NELSON, T. M., BARTLEY, S. H., & MACKAVEY, W. R. Responses to certain pseudo-isochromatic charts viewed in intermittent illuminance. *Percept. & Motor Skills*, 1961, **13**, 227-231.
15. NELSON, T. M., BARTLEY, S. H., & WISE, R. F. Size discrimination under two conditions of photic intermittency. *J. of Psychol.*, 1963, **56**, 219-225.
16. SENDERS, J. W., & SENDERS, V. L. Woodpile revisited: A comment on Gibbins and Howarths' failure to repeat Senders' findings. *Quart. J. Exper. Psychol.*, 1962, **14**, 176-177.
17. SENDERS, V. L. Visual resolution with periodically interrupted light. *J. Exper. Psychol.*, 1949, **39**, 453-465.
18. ———, On reading printed matter with interrupted light. *J. Exper. Psychol.*, 1954, **47**, 135-136.

*Department of Psychology*  
*Olds Hall*  
*Michigan State University*  
*East Lansing, Michigan 48823*



## PROBLEM-SOLVING MOTIVATION IN VERBAL CONDITIONING STUDIES\*<sup>1</sup>

*Department of Psychology, Pomona College*

---

ROGER E. VOGLER AND RUTH L. AULT

---

### A. INTRODUCTION

*A priori* assumptions are often made about the reinforcing functions of the contingent verbal stimuli (CVS) used to develop or maintain behavior. If the behavior changes as a function of the systematic presentation of the contingent stimuli, the investigator generally concludes that conditioning has occurred. If the investigator fails to suggest a variety of reinforcements, one is forced to assume that the investigator considers the CVS to be reinforcing (e.g., 7, 8).

Other investigators (e.g., 4, 6) have correlated personality factors with the degree to which an *S* "liked the reinforcers," but a more thorough examination is needed of the relationship between "liking the reinforcers," conditioning, and awareness of the response-reinforcement contingency. In the study reported here, it was hypothesized that problem-solving was a major source of reinforcement for a college population given a Taffel (9) type of verbal conditioning task. It was expected that this would be demonstrated in three ways: (a) many *Ss* would become aware of the response-reinforcement contingency (solve the problem), but would not be conditioned; (b) *Ss* would report problem-solving to be more reinforcing than the CVS in a postsession interview (PSI); and (c) the frequency of reporting problem-solving as a reinforcer would increase among *Ss* who were given a second problem to solve which was a reversal of the response class.

### B. METHOD

#### 1. Subjects

Twelve women and 25 men, enrolled in summer courses at the Claremont Graduate School, none of whom was in a psychology or education program, volunteered to be *Ss*.

---

\* Received in the Editorial Office, Provincetown, Massachusetts, on November 19, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This research was supported by a Pomona College Faculty Research Grant from funds made available through the Ford Foundation.

## 2. Apparatus

The apparatus and sentence construction tasks were modifications of those devised by Taffel (9). One hundred forty  $3 \times 5$  inch white index cards each had a different verb in the simple past tense typed on the center. Below the verbs were typed the four pronouns I, We, He, and They in random order, with each possible combination appearing with nearly equal frequency. The serial order of the cards was random, but each *S* received the same pattern.

A  $3 \times 5$  foot wooden screen was placed across the table, separating *S* from *E* to prevent nonverbal communication and to keep *S* from observing what *E* recorded. The screen was raised from the table top about  $1\frac{3}{4}$  inches to permit the presentation of the stimulus cards.

## 3. Procedure

*Ss* were told to make up one sentence for each index card presented to them, and to begin with one of the four pronouns and to use the verb on the card. Each sentence was spoken aloud. After each set of 10 cards (one trial block), *S* was handed a piece of paper, asked to write down any thought he had on the experiment (3), and then asked to return the paper to *E*.

The first trial block was used to calculate operant rate. During this time *E* remained silent. The CVS "Right" was assigned to whichever pair of pronouns—I and They or He and We—had the lower operant rate; "Wrong" was assigned to the higher rate pair. This was done so that each *S* would have an increased opportunity to raise his level of correct responding. Starting with the eleventh card, *E* said either "Right" or "Wrong" after every sentence spoken by *S*. This procedure is in accordance with the general technique of operant conditioning: base rate data are collected for a time interval prior to the introduction of the contingent stimuli. Moreover, this procedure fits the description of the 31 verbal conditioning studies reviewed by Krasner (5, p. 149), who observed "... the specific task given to the *S* is *not* presented as a learning one. The *S* is *not* asked to 'learn' a given series of words, nonsense syllables, or paired associates, nor is he even told that 'learning' is involved in the studies. Rather the *S* is asked to emit verbal behavior in terms of a given task, and the examiner attempts to reinforce a preselected class of his 'verbal behavior' by carefully controlled verbal or nonverbal behavioral cues."

If an *S* had not written an hypothesis in his "thoughts about the experiment" that one or more pronouns were right or wrong, by the 80th card (after eight trial blocks), the learning session was terminated. Of those *Ss* who had written a correct contingency hypothesis, half were terminated six trial



blocks after the statement was written. The other half was given two trial blocks past the awareness statement and then the contingency was reversed. If the Ss wrote a second correct contingency hypothesis, they were given two trial blocks past the second awareness statement and then terminated. If they did not report awareness of the contingency reversal within four trial blocks, they were terminated. Ss judged aware during the learning session were alternately assigned to the reversal and nonreversal groups.

Following the learning session, all Ss were asked seven questions (the PSI). These questions were designed to reassess Ss' awareness of the correct contingency, to investigate Ss' goals during the task, and to determine Ss' own evaluation of the reinforcing aspects of the task and of the CVS.

### C. RESULTS

. An examination of the correct response rate (i.e., the rate of the pronouns upon which "Right" was contingent) revealed no differences between aware and unaware Ss, nor between aware-reversed and aware-nonreversed Ss<sup>2</sup> both when aligned for the trial block of awareness and when not aligned. Differences in the average percentage correct response rate between preawareness and postawareness for individual Ss ranged from -17 to +48 percentage points. Four Ss decreased their correct response rates (-17, -10, -10, and -3 percentage points), four increased by at least 20 percentage points (+22, +33, +44, and +48), while the remaining 16 aware Ss increased by less than 15 percentage points. Thus there was little evidence of conditioning for Ss who were aware of the contingency between "Right" and the contingent pronouns.

All Ss were asked in the PSI whether they had tried to figure out why or when the CVS were presented. Their responses fell into three categories: yes, no, and only at the beginning. Only one of 37 Ss did not try to solve the problem and one-third of the unaware Ss did not continue the attempt for very long (see Table 1).

Two PSI questions were designed to investigate approval-seeking and disapproval-avoidance behavior. Question 4 read, "Did you try to say sentences which made the experimenter say 'Right,' 'Wrong,' both, or neither?" The choice "Right" was interpreted as approval-seeking, the answer "Wrong" as problem-solving methods or other reasons, "both" was considered problem-solving, while "neither" indicated indifference or lack of positive reinforcement.

---

<sup>2</sup> Three Ss were not judged aware until the PSI and were thus assigned to the aware-nonreversed group.

ment of the CVS. Most *Ss* were indifferent to the CVS or were trying to solve the problem (see Table 2). Less than one-fourth of the *Ss* worked to hear "Right."

TABLE 1  
INSTANCES OF PROBLEM-SOLVING BEHAVIOR BY GROUPS

Group	(N)	Yes	No	Beginning only
Unaware	(15)	9	1	5
Aware-nonreversed	(13)	13	0	0
Aware-reversed	(9)	8	0	1
Totals	(37)	30	1	6

TABLE 2  
RESPONSE CATEGORY AFFIRMED BY *Ss* IN EACH GROUP TO QUESTION 4

Group	(N)	"Right"	"Wrong"	Both	Neither
Unaware	(15)	2	0	1	12
Aware-nonreversed	(13)	4	1	6	2
Aware-reversed	(9)	2	1	4	2
Totals	(37)	8	2	11	16

Similarly, Question 5 read, "Did you try to say sentences which let you avoid hearing *E* say 'Right,' 'Wrong,' both, or neither?" As Table 3 indicates, most *Ss* found neither CVS aversive.

TABLE 3  
RESPONSE CATEGORY AFFIRMED BY *Ss* IN EACH GROUP TO QUESTION 5

Group	(N)	"Right"	"Wrong"	Both	Neither
Unaware	(15)	1	0	0	14
Aware-nonreversed	(13)	1	1	1	10
Aware-reversed	(9)	0	2	0	7
Totals	(37)	2	3	1	31

In further substantiation of the hypothesis that the major source of reinforcement for *Ss* was problem-solving behavior and not the presence or absence of the CVS, *Ss* were asked to rank from one to four (one being the highest) the four "goals" they had while the experiment was in progress. Ten alternatives were listed, including solving the problem, getting *E*'s approval by hearing "Right," avoiding *E*'s disapproval by not hearing "Wrong," and an "other" category which provided an opportunity to specify choices not listed. Table 4 indicates the percentage of *Ss* in each group who listed the specified alternative in one of the first four positions.

Alternative (C), solving the problem, received the most first choices and also the most first and second choices combined (see Table 5), while alternative (A), saying the first thing that came to mind, was the second most frequent first and combined choice.

TABLE 4  
PERCENTAGE OF *Ss* RANKING AN ALTERNATIVE IN ONE OF THE FIRST FOUR POSITIONS

Selected alternatives <sup>a</sup>	Unaware	Aware- nonreversed	Aware- reversed
(C) Solving the problem	60%	61%	100%
(A) Saying the first thing that comes to mind	86%	69%	77%
(H) Purposely varying the order of the pronouns	67%	61%	67%
(D) Wanting <i>E</i> 's approval	26%	53%	33%
(I) Did not care about the pronoun choice—random	67%	15%	33%
(B) Telling a story	53%	61%	22%

<sup>a</sup> All other alternatives were chosen by less than 50 per cent of each group.

TABLE 5  
NUMBER OF *Ss* RANKING AN ALTERNATIVE AS FIRST OR SECOND CHOICE

Alternative	Choices		
	1st	2nd	1st and 2nd combined
(C)	17	5	22
(A)	10	10	20
(H)	2	6	8
(D)	1	6	7

The four *Ss* who increased their response rate by at least 20 percentage points—i.e., the four whose response rate indicated conditioning—all selected (C) as their first choice. Three of four chose (D), wanting *E*'s approval, for their second choice, and the other chose (H), purposely varying the pronoun order.

#### D. DISCUSSION

Only four *Ss*, all aware, were conditioned; yet 18 other aware *Ss* failed to be conditioned. No unaware *Ss* were conditioned, and no significant differences were obtained between the conditioning rates of aware and unaware groups of *Ss*, nor between aware-reversed and aware-nonreversed groups. These data are not in accordance with the finding of other workers that *Ss* are conditioned when "Right" and "Wrong" are used as the CVS (e.g., 2, 8).

One possible reason for the failure to obtain conditioning as have others

may involve the extent to which nonverbal communication factors contribute to the conditioning process. No other study reports that *Ss* said anything on first hearing the CVS. In this study, 22 of 37 *Ss* questioned *E* on first hearing the CVS, and four more did so the second time a CVS was presented. *S*'s queries showed confusion, surprise, or amusement. To questions such as "Pardon me?" *E* repeated the CVS by saying, "I said ———." If *S* asked for more specific information, such as "How can anything be 'Wrong'? I gave you a complete sentence," *E* replied "Please go on with the experiment." These particular responses were chosen by *E* for three reasons: (*a*) to avoid inducing a problem-solving set—as might occur with "It's part of the experiment" or "You have to figure that out"; (*b*) to convey as little additional information as possible, so that nonquestioning *Ss* would not be at a disadvantage; and (*c*) to discourage further questioning by *S*, while at the same time not totally ignoring the question.

It is probable that much nonverbal information would have been communicated if *S* and *E* were face to face. However, unlike other studies—with the exception of Ault and Vogler (1)—a screen was placed between *S* and *E* in this study to prevent nonverbal communication. Yet in none of the studies in which *Ss* were conditioned is it suggested that nonverbal communication was a factor. It is quite possible, if not probable, that when *S* and *E* are face to face the many nonverbal discriminative and reinforcing stimuli presented concomitantly with "Right," "Wrong," "Good," "Hmhm," etc. may contribute significantly to the conditioning of *Ss*' verbal behavior.

Since *Ss* who were fully aware of the contingency between their behavior and the "Right" or "Wrong" spoken by *E* did *not* show a significant increase in correct response rate, it is concluded that "Right" possessed little positive reinforcing power for this subject population and that "Wrong" was not aversive enough to result in avoidance behavior from *Ss*. Inasmuch as the PSI data indicate clearly that problem-solving was the primary reinforcer for most *Ss*, and especially so for aware *Ss*, it is more reasonable to infer that the CVS served primarily as discriminative stimuli for the reinforcer of problem-solving. That problem-solving is the major reinforcer is also supported by the fact that *Ss* given an additional opportunity to earn it attempted to do so and reported it to be more important than did other *Ss*.

An alternative explanation to the more frequent ranking of problem-solving by the aware-reversed group is that they spent more time at it than did the other groups. However, one can infer from *Ss*' problem-solving behavior that this was a reinforcing task for them: even though *Ss* were not instructed to find the contingency, they attempted to do so. In fact, several *Ss*, after writing

a correct contingency statement, continued to seek another, usually more complex, solution as judged by their subsequent highly variable response rates, their "thoughts about the experiment," and the PSI data. Ss appeared to think the problem more difficult, and rather than settle for attaining one hundred per cent "reinforcement" in the form of hearing "Right," they continued to problem-solve (and to hear "Wrong").

### E. SUMMARY

Data indicate that problem-solving behavior was the major source of reinforcement for college students in a verbal conditioning task and that the contingent verbal stimuli, contrary to assumptions often made *a priori*, were only mildly reinforcing. Only four Ss, all aware, were conditioned; yet 18 other aware Ss failed to be conditioned. No unaware Ss were conditioned, and no significant differences were obtained between the conditioning rates of aware and unaware groups of Ss. Factors involved in the failure here to obtain the conditioning which has been reported by others are (a) the use of a screen between S and E to prevent nonverbal discriminative and reinforcing stimuli from contributing to the conditioning process, and (b) differences in S population.

### REFERENCES

1. AULT, R. L., & VOGLER, R. E. Discriminative and reinforcing functions of four verbal stimuli. *Psychol. Rep.*, 1968, in press.
2. BUCHWALD, A. M. Extinction after acquisition under different verbal reinforcement combinations. *J. Exper. Psychol.*, 1959, **57**, 43-48.
3. DENIKE, L. D. The temporal relationship between awareness and performance in verbal conditioning. *J. Exper. Psychol.*, 1964, **68**, 521-529.
4. DOHERTY, M., & WALKER, R. The relationship of personality characteristics, awareness, and attitude in a verbal conditioning situation. *J. Personal.*, 1966, **34**, 504-515.
5. KRASNER, L. Studies of the conditioning of verbal behavior. *Psychol. Bull.*, 1958, **55**, 148-170.
6. LEVIN, S. M. The effects of awareness on verbal conditioning. *J. Exper. Psychol.*, 1961, **61**, 409-416.
7. LEVINE, M., LEITENBERG, H., & RICHTER, M. The blank trials law: The equivalence of positive reinforcement and nonreinforcement. *Psychol. Rev.*, 1964, **71**, 94-103.
8. SPENCE, J. T. Effects of verbal reinforcement combination and instructional condition on performance of a problem-solving task. *J. Personal. & Soc. Psychol.*, 1966, **3**, 163-170.
9. TAFFEL, C. Anxiety and the conditioning of verbal behavior. *J. Abn. & Soc. Psychol.*, 1955, **51**, 496-501.

Department of Psychology  
Mason Hall  
Pomona College  
Claremont, California 91711





## FACTORIAL STRUCTURE OF CQT, ACT, AND SAT TEST SCORES FOR 50 AVAILABLE COLLEGE FRESHMEN\*

*Department of Educational Psychology, University of Wisconsin-Milwaukee*

---

RUSSELL N. CASSEL AND ARDEN C. EICHSTEADT

---

### A. INTRODUCTION

This study was concerned with making a comparative analysis of the factorial structure for the College Qualification Tests (CQT), the American College Tests (ACT), and the Scholastic Aptitude Tests (SAT) for 50 available freshmen participants at the University of Wisconsin-Milwaukee (UWM) for the 1967-1968 school year. It seeks specifically to ascertain the test content structure of the CQT, ACT, and SAT test batteries, and to identify the factorial base used as a means for predicting scholastic success of entering college students by each of the separate batteries.

### B. METHOD

The 50 available college freshmen were selected from among the approximately 3,000 entering freshmen students in September, 1967, at UWM. For these participants scores on all three test batteries were available: i.e., CQT, ACT, and SAT. With but a few exceptions, the participants of this study were graduates from local high schools (Wisconsin) in the Spring of 1967.

The ACT and SAT test batteries were administered in the respective local high schools of participating members during the high school senior year, and in cooperation with the particular sponsoring agency for the tests: i.e., the ACT by the American College Testing Program, Iowa City, Iowa; and the SAT by the Educational Testing Service, Princeton, New Jersey. The CQT tests are a portion of the UWM regular college testing program, and they were administered at UWM prior to student registration.

### C. DISCUSSION

Both the SAT and the CQT tests incorporate the traditional test philosophy that certain select verbal and quantitative behavior samples made independent of each other serve as the best generalized predictors of college scho-

---

\* Received in the Editorial Office, Provincetown, Massachusetts, on November 20, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

lastic success. The CQT battery has added a third element in a separate test dealing with "information," but with no change in basic testing philosophy. Thus, from the SAT and CQT tests we have the notion that scholastic aptitude represents a generalized phenomena that is equally applicable in a wide diversity of academic specialized areas, and can be best assessed in the form of "capability" or potential of the individual.

The ACT, to be sure, has introduced a new and different basic test philosophy for use in predicting college scholastic success. Germane to the new philosophy is the idea that demonstrated ability in "hard core" scholastic areas (mathematics, science, social studies, and language arts) represents the best predictor of college scholastic success. The nucleus of this concept derives from the notion that the best predictor of success in mathematics is a test of present level of achievement in mathematics; the best predictor of success in science is a test of present level of achievement in science; the best predictor of success in social studies is a test of present level of achievement in social studies; and similarly for language arts. Thus, the ACT test philosophy eliminates the capability concept representing the potential of the individual, and substitutes ability representing school success; and broadens the factorial test base from verbal and quantitative data to the four "hard core" curriculum areas.

#### D. RESULTS

Table 1 contains the intercorrelation matrix for all of the variables of the three test batteries used in the study: i.e., CQT, ACT, and SAT. These correlations served as the basis for factors extracted in Tables 2 and 3.

Six independently organized factors were extracted, each with significant factor loadings, through the use of the principal component method for the CQT, ACT, and SAT tests in this study, and as contained in Table 2. These factors have been rotated to simple structure through the use of the varimax-orthogonal method, and as contained in Table 3. From the factor loadings and test identifications, these six factors are named as follows: I-verbal, II-quantitative, III-English, IV-natural science, V-social studies, and VI-information.

From Table 4 it may be observed that all three test batteries (CQT, ACT, and SAT) employ the traditional verbal and quantitative factorial base for prediction purposes. There is good agreement among the three test batteries as to what the verbal and quantitative base is, but the better agreement is in the quantitative area. The CQT tests reflect a somewhat purer form of the verbal component than the other two test batteries, and the SAT reflects a somewhat less pure form of the quantitative factor than the other two.

TABLE 1  
INTERCORRELATION MATRIX  
( $N = 50$ )

Variables	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$	$X_7$	$X_8$	$X_9$
1. CQT- $X_1$ Verbal	•1000	124	528	540	074	664	544	830	153
2. CQT- $X_2$ Numerical		1000	324	083	839	214	359	092	701
3. CQT- $X_3$ Information			1000	155	232	509	450	557	326
4. ACT- $X_4$ English				1000	069	429	365	588	129
5. ACT- $X_5$ Mathematics					1000	209	339	049	776
6. ACT- $X_6$ Soc. Studies						1000	699	729	350
7. ACT- $X_7$ Nat. Science							1000	583	441
8. SAT- $X_8$ Verbal								1000	200
9. SAT- $X_9$ Mathematics									1000
Mean	52.58	35.68	51.54	22.02	24.76	24.40	25.38	515.80	526.60
SD	11.31	10.34	12.25	3.86	5.88	4.83	4.83	83.07	81.64

• Decimal points eliminated.

TABLE 2  
PRINCIPAL COMPONENT FACTOR ANALYSIS  
( $N = 50$ )

Variables	I	II	III	IV	V	VI
1. CQT-X <sub>1</sub> Verbal	*770	-453	-013	-168	003	-336
2. CQT-X <sub>2</sub> Numerical	541	738	051	-165	-209	-140
3. CQT-X <sub>3</sub> Information	680	-055	-587	-333	-078	260
4. ACT-X <sub>4</sub> English	553	-368	662	-222	-078	243
5. ACT-X <sub>5</sub> Mathematics	512	789	130	-064	004	-117
6. ACT-X <sub>6</sub> Soc. Studies	-823	-249	-078	309	197	-021
7. ACT-X <sub>7</sub> Nat. Science	802	-031	-031	473	-315	087
8. SAT-X <sub>8</sub> Verbal	802	-477	-003	-093	117	-085
9. SAT-X <sub>9</sub> Mathematics	611	652	067	067	330	160

\* Decimal points eliminated.

TABLE 3  
FACTORS ORTHOGONALLY ROTATED  
( $N = 50$ )

Variables	I	II	III	IV	V	VI	h <sup>2</sup>	Factor ident.
1. CQT-X <sub>1</sub> Verbal	*902	052	219	157	-115	192	940	I
2. CQT-X <sub>2</sub> Numerical	042	932	020	119	-155	147	931	II
3. CQT-X <sub>3</sub> Information	357	185	-001	161	042	898	995	VI
4. ACT-X <sub>4</sub> English	353	037	923	126	021	-003	994	III
5. ACT-X <sub>5</sub> Mathematics	026	952	007	090	064	018	919	II
6. ACT-X <sub>6</sub> Soc. Studies	668	114	101	519	341	153	880	IV
7. ACT-X <sub>7</sub> Nat. Science	346	261	146	860	015	160	975	IV
8. SAT-X <sub>8</sub> Verbal	814	-006	314	231	112	269	900	I
9. SAT-X <sub>9</sub> Mathematics	070	802	060	162	498	127	942	II

\* Decimal points eliminated.



TABLE 4  
COMPARISON OF FACTORIAL CONTENT OF CQT, ACT, AND SAT TESTS  
( $N = 50$ )

Factor identification	CQT Tests	ACT Tests	SAT Tests
I-Verbal	.902	.668	.814
II-Quantitative	.932	.952	.802
III-English		.923	
IV-Natural Science		.860	
V-Social Studies		.341	.498
VI-Information	.898		

The CQT battery has indeed added a factorial component through the information test and, thus, has broadened the prediction factorial base. Whether this broadened base has increased prediction success is not pertinent to this study.

The ACT battery has similarly broadened its factorial base for prediction purposes by the English and natural science factors, and by the marginally significant loading on the social studies factor. Thus, it becomes obvious that, for this study, the ACT battery has a much broader factorial base for predicting college scholastic success than do either the SAT or CQT test batteries. Again, whether or not this broadened base increases prediction accuracy is not pertinent to this study.

#### E. SUMMARY

A comparison was made of the factorial structure of the CQT, ACT, and SAT tests for scores on 50 available college freshmen at the University of Wisconsin-Milwaukee through the use of the principal component factorial method. Each of the six factors demonstrated independence of organization evidenced by significant factor loadings, and was rotated to simple structure through the use of the varimax-orthogonal procedure. Based on factor loadings and test names for the variables with significant loadings, the factors were identified as follows: I-verbal, II-quantitative, III-English, IV-natural science, V-social studies, and VI-information. All three of the test batteries (CQT, ACT, and SAT) used the traditional verbal and quantitative factor base. The CQT displayed an additional information factor base. The ACT battery reflected a considerably broader factorial base, which included three additional factors representing the "hard core" curriculum areas: English, natural science, and social studies.

#### REFERENCES

1. BENNETT, G. K., *et al.* College Qualification Tests Manual. New York: Psychological Corp., 1957.

2. FORTNA, R. O. A factor-analytic study of the Cooperative School and College Ability Tests and Sequential Tests of Educational Progress. *J. Exper. Educ.*, 1963, 32, 1187-1190.
3. LINDQUIST, E. F. An evaluation of a technique for scaling high school grades to improve prediction of college success. *Educ. & Psychol. Meas.*, 1963, 23, 623-646.

*Department of Educational Psychology*  
*School of Education*  
*The University of Wisconsin-Milwaukee*  
*Milwaukee, Wisconsin 53201*



CHOICES FOR PLAYER 1	CHOICES FOR PLAYER 2	
	$C_2$	$D_2$
$C_1$	$X_1, X_1$	$X_2, X_3$
$D_1$	$X_3, X_2$	$X_4, X_4$

FIGURE 1  
GENERAL FORM OF PRISONER'S DILEMMA PAYOFF MATRIX

repeated-play situation, since they consistently achieve outcomes more favorable than those accruing to rational strategy. Rapoport and Chammah (8) summarize this point:

Ordinary mortals . . . when playing Prisoner's Dilemma many times in succession hardly ever play (the dominant) strategy one hundred percent of the time. Evidently the run-of-the-mill players are not strategically sophisticated enough to have figured out that (the dominant) strategy is the only rationally defensible strategy, and this intellectual shortcoming saves them from losing.

Shure and Meeker (13) showed that under the limiting condition of repeated play the average college student *S* does, in fact, adopt a "rationally" prescribed strategy in which he persistently chooses the dominant alternative and strategically ignores the choices of the other player.

Bieri (1, 2), using an adaptation of Kelly's Role Repertory Test (3), found some college student *Ss* to be relatively high in "cognitive complexity": their discovered systems of constructs differentiate highly people in the environment. A construct system which provides poor differentiation among persons is considered to be cognitively simple in structure.

Larsen (4), using Bieri's (1) "Repertory test" as the measure of cognitive complexity and Rokeach's (10) dogmatism scale, found that those subjects who were high in cognitive complexity and low in dogmatism also placed low on the chauvinism-nationalism scale and were more able than other subjects to sustain cooperative relationships in mutually dependent game situations.

The experimental design of the present authors aimed at administering Bieri's paper and pencil test to approximately 200 college students, by selecting out randomly 24 students who were at the upper and lower 10th per-

tiles in scores, and by replicating the Shure and Meeker-prisoner's dilemma game conditions.

It was the hypothesis of the present authors that those students highest in cognitive complexity would also possess the empathy or the strategic sophistication to choose the nondominant move, to be more capable than the other students of sustained cooperativeness in game behavior.

The purpose of the present study was to determine the independent and interaction effects of the variable "cognitive complexity" on the behavior of Ss playing a series of modified prisoner's dilemma games.

In the present study a modification of the prisoner's dilemma game was employed. In the classical game, the payoff to Player 1, when both players make a competitive response, was greater than the payoff to Player 1 when he made a cooperative and Player 2 a competitive response. For this experiment, the condition was reversed. The payoff to Player 1 was less when both he and Player 2 competed than when he cooperated and Player 2 competed. Through this modification, an attempt was made to increase the utility of making cooperative responses. In addition, in the present game, the subject, unknowingly, played against a simulated other player who responded according to a fixed schedule of cooperative and competitive choices. In most game experiments there are two subjects who may freely vary their choices in an attempt to modify each other's behavior. The design of this present study did not permit this form of reciprocal bargaining. It did, however, provide a relatively constant social stimulus for the subject: i.e., an "other" player who predominantly emitted cooperative or competitive responses.

The variable "high in cognitive complexity" or "low in cognitive complexity" could be identified as a personality dimension. It is felt that persons with highly differentiated constructs, those "high in cognitive complexity," would be more likely to select the cooperative alternative than those who were simplistic in orientation—black and white thinkers. The second variable, the strategy of the other player, could be classified as a situation dimension. Shure and Meeker and others contend that, in repeated play PD, subjects do not make more cooperative responses to an 85 per cent cooperative response strategy by the other player than to an 85 per cent competitive response strategy. The final variable, subjects' responses across 20 trial blocks, was included because of the previous research findings which indicated that the proportion of cooperative responses made by the subjects decreases through time.

The major hypotheses of the present study were independent variable effects and interaction effects, under independent variable effects: those high in



cognitive complexity would make significantly more cooperative responses than the lows; the strategy of the other player would not significantly affect the number of cooperative responses emitted by the subjects; the number of cooperative responses made by subjects will decrease significantly through time and under interaction effects, those "high in cognitive complexity" would make significantly more cooperative responses to the cooperative other strategy than to the 50-50 or competitive other strategies; Ss "low in cognitive complexity" would tend to take advantage of a cooperative other strategy, and thus would make significantly fewer responses to the cooperative other strategy than to the 50-50 or competitive other strategies.

The interaction effects hypothesis represented an attempt to account for previous findings that indicate that systematic variations in the cooperativeness or competitiveness of the strategy of the other player did not affect the subject's choices. In essence, the authors are hypothesizing that subjects who are high and low in cognitive complexity respectively respond differently to a cooperative strategy on the part of the other player. Thus any differences between responses across all subjects to the strategy of the other player are, in effect, cancelled out by the systematic differences between those high and low on the test of cognitive complexity.

### B. METHOD

The Prisoner's Dilemma variation, as employed in this study, appears in Figure 2. In this game, Player 1 controlled the columns with A representing the cooperative alternative and B the competitive, while Player 2 controlled the columns with X representing the cooperative alternative and Y the competitive. The outcomes or payoffs which resulted from the mutual choice of both players appear in each of the four cells of the matrix.

		PLAYER 1 Choices	
		A	B
PLAYER 2 CHOICES	X	Player 1: +3¢ Player 2: +3¢	Player 1: +5¢ Player 2: 0¢
	Y	Player 1: 0¢ Player 2: +5¢	Player 1: -3¢ Player 2: -3¢

FOUR CELLS

FIGURE 2  
PAYOFF MATRIX—THE PRISONER'S DILEMMA GAME

### 1. Subjects

An 18-item Test of Cognitive Complexity adapted by Bieri (1) from Kelly's Role Repertory Test (3) was administered to approximately 200 students enrolled in introductory classes at California State College at Los Angeles. This scale defines the cognitively complex individual as one who displays a high degree of differentiation in concepts towards the significant others in his environment. The S "low in cognitive complexity" is one who is very simple in conceptual structure, one who makes few differentiations among significant others. Twenty-four students who scored in the upper decile of the Bieri cognitive complexity scale and 24 who scored in the lowest decile were selected as subjects for the present study. An equal number of subjects of both groups was assigned to one of three treatment conditions (*cf.* Table 1).

TABLE 1  
MEAN NUMBER OF COOPERATIVE CHOICES BY 20 TRIAL BLOCKS FOR HIGH COGNITIVE COMPLEXITY AND LOW COGNITIVE COMPLEXITY UNDER COOPERATIVE, 50-50, COMPETITIVE TREATMENTS

Personality dimension	Treatments								
	Cooperative			50-50			Competitive		
	Trial blocks			Trial blocks			Trial blocks		
	First	Second	Third	First	Second	Third	First	Second	Third
High cognitive complexity	10.00	8.88	8.88	9.75	7.12	9.12	8.38	6.62	8.62
Low cognitive complexity	8.88	3.75	2.38	9.12	7.50	6.38	9.12	5.38	7.00

### 2. Experimental Procedure

Either two subjects, or a subject and a confederate were seated opposite one another. A partition obstructed their view of each other. The subject was provided with a game board with an indicator light located in each of the four payoff quadrants: i.e., AX, AY, BX, and BY. At the bottom of each game board were switches which enabled Player 1 to choose Column A or B, and Player 2, Row X or Y.

Instructions stressing neither a cooperative nor a competitive orientation to the game were read to the subjects who were given \$2.25 with which to start the game. Ss were informed that they could retain the amount of money which they possessed when the game was completed, but were informed of trials they played. They were also given individual scoreboards on which to keep track of their individual wins and losses.

After the experimenter left the room, 60 consecutive trials were run. The experimenter, unbeknown to the subjects, assumed the role of the "other" player for both subjects. A predetermined program of experimenter controlled "other" player responses was used for each of the treatment groups: cooperative simulated other—the simulated other made 85 per cent cooperative and 15 per cent competitive choices; 50-50 simulated other—the simulated other played a 50 per cent cooperative, 50 per cent competitive strategy; competitive simulated other—the simulated other player made 15 per cent cooperative and 85 per cent competitive choices. For the 85 per cent cooperative and the 50-50 treatment groups, the appropriate proportions of cooperative "responses" by the simulated other were located randomly within the 60-trial sequence. For the 85 per cent competitive treatment, the cooperative schedule was reversed.

Each trial lasted 20 seconds during which time the subject and the "other" player made a competitive or a cooperative choice: a light indicated which cell in the payoff matrix should have been used to determine wins and losses; the subject recorded his own wins and losses; and the examiner recorded the subject's choice. At the end of 60 trials the subjects were questioned to determine whether they had perceived the experimenter's intervention. They were then informed of the experimenter's intervention, were requested not to discuss it, and were awarded the original sum of money with which they had begun the game.

### C. RESULTS

Table 1 shows the mean number of cooperative responses made by Ss high *versus* low in cognitive complexity under the three strategy conditions. It is possible (7, 9) that the S's performance might change as he became familiar with the game and learned the strategy that the other player was utilizing. Since this might have been the case, the results were set out in three blocks of 20 trials each. A mixed design analysis of variance was employed, and blocks of trials were used as a third independent variable. The results of this analysis are shown in Table 2.

In the between-subjects part of Table 2, the only comparison that reached significance was the difference between the high and the low cognitively complex, with the highs emitting more cooperative responses than the lows. Two effects were significant in the within-subjects part of the analysis, blocks of trials, and the interaction between high-low and blocks of trials. The linear and quadratic components were extracted from the trials effect, and both were significant at the .005 level. The linear component indicates that there was a significant tendency for all subjects to make fewer cooperative responses as the

TABLE 2  
ANALYSIS OF VARIANCE  
Personality Dimension  $\times$  Strategy  $\times$  Trial Blocks

Source	df	MS	F
Between subjects			
Personality (A)	1	142.0	4.38*
Strategy (B)	2	13.27	.41
A $\times$ B	2	46.39	1.43
Error (b)	42	32.29	
Total	47		
Within subjects			
Trial Blocks (C)	2	95.90	11.57**
A $\times$ C	2	32.50	3.92*
B $\times$ C	4	10.79	1.30
A $\times$ B $\times$ C	4	7.09	.93
Error (w)	84	8.29	
Total	96		

\*  $p < .05$ .

\*\*  $p < .005$ .

experiment progressed. However, the quadratic component indicates that the deviation from linearity was also significant, showing that the subjects were returning to a more cooperative strategy in the last block of 20 trials.

Inspection of the significant interaction between high and low complexity groups and trials shown in Figure 3 indicated that this return to a more cooperative strategy was demonstrated by the highs only. Although both groups of subjects became more competitive as the game went on, the lows became significantly more competitive than the highs. It should be noted, however, that for neither group at any stage of the experiment did cooperative responses reach the 50 per cent level.

The analysis further indicates that the type of strategy employed against the subject had no significant effect upon his behavior, either directly or in interaction with the complexity differences or the trials effect.

#### D. DISCUSSION

This study demonstrates that those Ss high in cognitive complexity make more cooperative choices than do the lows. Actually, the findings suggest that the highs make fewer competitive responses, since at no time did the mean number of cooperative responses for either group exceed the 50 per cent level even though the highs did approximate this level. The data also indicate a diminishing number of cooperative responses over time, similar to the effect found by Scodel *et al.* (12) and Minas *et al.* (6).

Of particular interest is the finding that the strategy employed by the sim-

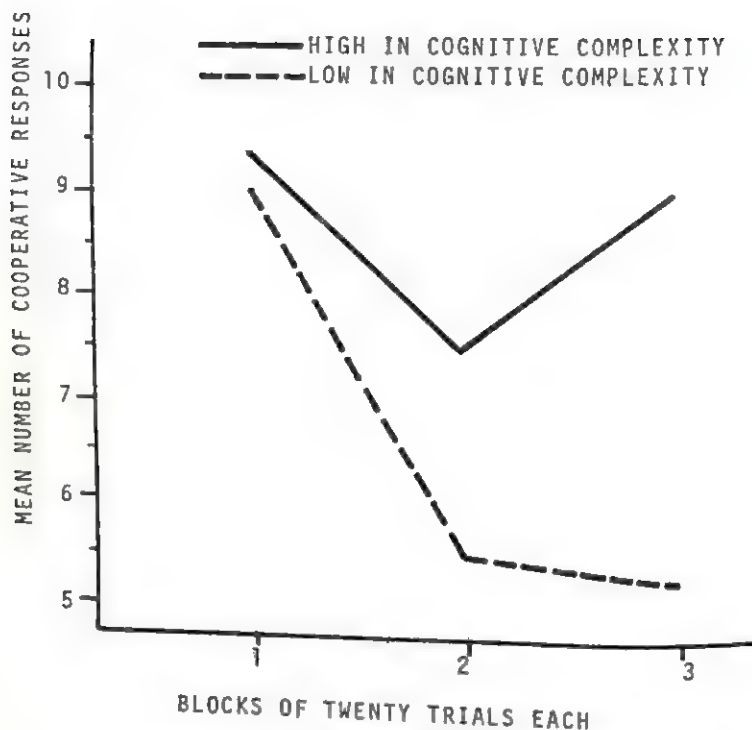


FIGURE 3  
MEAN NUMBER OF COOPERATIVE CHOICES BY 20 TRIAL BLOCKS  
FOR Ss HIGH AND LOW IN COGNITIVE COMPLEXITY

ulated other had no overall or differential effects upon the behavior of the highs or the lows. This finding led to a rejection of the interaction effect hypothesis. It might be possible that no effects were observed because the Ss were not aware of the differences in the other person's strategies. However postsession questionnaires indicated that both the highs in cognitive complexity and the lows attributed different characteristics to their "opponent" and to the task as a function of the strategy condition to which they had been subjected.

Two possible interpretations of the significant differences between the highs and lows seem possible. The first interpretation involves analyzing the consequences of different player strategies for his own and the other player's outcomes. For instance, when faced with an 85 per cent cooperative strategy, a more competitive strategy on the part of the player increases his own gains at the expense of his opponent's gains. A less competitive strategy reduces the player's gains somewhat, while increasing his opponent's gains. In the 50-50 and 85 per cent competitive conditions, there is no strategy that the player can



develop which will provide large gains for himself. His choice of strategy can, however, control, to a large extent, his opponent's outcomes. A more competitive strategy in these conditions decreased the subject's gains or increased his losses, and also increased his opponent's losses. A less competitive strategy decreased the subject's losses, but increased his opponent's gains.

The fact that those Ss low in cognitive complexity made more competitive responses under all three of these conditions would seem to indicate that, when the opportunity to increase their own gains occurred, they utilized it, even though that meant decreasing the other player's gains. When they could not win large sums they were more likely to use a strategy that increased their opponent's losses, even though that meant taking a loss themselves. Those Ss high in cognitive complexity were more likely to use a strategy that resulted in a more equitable distribution of payoffs, if their opponent was cooperative, or one that minimized their own losses if he was not. Unfortunately, because of the values of the payoff matrix, it was not always possible to determine whether the player's own outcomes or his opponent's outcomes were providing the motivation to use a particular strategy.

A second interpretation of these findings could be derived from the assumption that Ss low in cognitive complexity have a generalized tendency to play more competitively than the highs, regardless of the outcomes for themselves and the other player. Such a behavioral predisposition would seem to follow from the personality characteristics one might postulate to be associated with the highs and the lows. Such an assumption would explain in part the significant interaction in the analysis of variance (Table 2) which indicates that the lows played more competitively as the game progressed, whereas the highs tended to become more competitive in the middle of the game, but returned almost to their original, less competitive style at the end of the games.

The data do not provide a basis for stating unequivocally whether one or both of these interpretations is correct: i.e., whether subjects select a response strategy based on projected outcomes for self and other, or a generalized tendency to compete or cooperate regardless of the outcomes. The present study substantiates that an individual's style of life in terms of tendency to be relatively *more* or *less* complex is related to his behavior in two-person games where his own and the other player's outcomes are affected by his making cooperative-competitive choices.

#### E. SUMMARY

Of 200 college student Ss, 24 were selected out randomly who were at the upper and lower 10th percentiles on Bieri's test of cognitive complexity. Those

Ss high in cognitive complexity demonstrated highly differentiated constructs. Both groups played against *E* in the Prisoner's Dilemma game situation, ostensibly playing against another *S*. In the overall, Ss high in cognitive complexity emitted more cooperative responses than the lows though there was a significant tendency for all Ss to make fewer cooperative responses as the experiment progressed.

The type of strategy employed against the Ss had no significant effect on their behavior either directly or in interaction with the complexity differences or the trials effect. The strategy employed by the simulated other had no overall or differential effect on the highs or the lows. It is possible that no effects were observed because Ss were not aware of the differences in the other person's strategies.

The fact that Ss low in cognitive complexity made more competitive responses under all conditions would seem to indicate that, when the opportunity to increase their own gain occurred, they utilized it even when that meant decreasing the other person's gains. Those Ss high in cognitive complexity were more likely to use a strategy that resulted in a more equitable distribution of payoffs if their opponent was cooperative, or one that minimized their own losses if he was not. *S*'s style of life, in terms of tendency to be relatively more or less complex, is related to his behavior in two-person games where his own and the other person's outcomes are affected by his making cooperative-competitive choices.

#### REFERENCES

1. BIERI, J. Cognitive complexity, simplicity and predictive behavior. *J. Abn. & Soc. Psychol.*, 1955, 51, 263-268.
2. ———. Complexity-simplicity as a personality variable in cognitive and preferential behavior. In D. W. Fiske & R. M. Maddi (Eds.), *Functions of Varied Experience*. Homewood, Ill.: Dorsey Press, 1961.
3. KELLY, G. A. The Psychology of Personal Constructs (Vol. 1). A Theory of Personality. New York: Horton, 1955.
4. LARSEN, K. S. Cognitive complexity and dogmatism as predictors of chauvinistic nationalism and game behavior. Unpublished Master's thesis, California State College at Los Angeles, 1966.
5. LUCE, R. D., & RAIFFA, H. Games and Decisions. New York: Wiley, 1957.
6. MINAS, J. S., SCODEL, A., MARLOWE, D., & RAWSON, H. Some descriptive aspects of two person, non-zero-sum games: Part II. *J. Conflict Resolut.*, 1960, 4, 193-197.
7. RAPOPORT, A. Optimal policies for the prisoner's dilemma. *Psychol. Rev.*, 1967, 74, 136-138.
8. RAPOPORT, A., & CHAMMAH, A. M. Prisoner's Dilemma. Ann Arbor, Mich.: Univ. Michigan Press, 1965.
9. RAPOPORT, A., & ORWANT, C. Experimental games: A review. *Behav. Sci.*, 1962, 7, 1-37.

10. ROKEACH, M. *The Open and Closed Mind*. New York: Basic Books, 1960.
11. SCHELLING, T. C. *The Strategy of Conflict*. Cambridge, Mass.: Harvard Univ. Press, 1960.
12. SCODEL, A., MINAS, J. S., RATOOSH, P., & LIPETZ, M. Some descriptive aspects of two person, non-zero-sum games. *J. Conflict Resolut.*, 1959, 3, 114-119.
13. SHURE, G. H., & MEEKER, R. J. Empirical demonstration of normative behavior in the prisoner's dilemma. *Proc. 76th Annual Convention, American Psychological Association*, Washington, D. C., 1968, 61-62.

*Department of Psychology*  
*California State College at Los Angeles*  
*5151 State College Drive*  
*Los Angeles, California 90032*



## RELATIONSHIPS BETWEEN EXPRESSED PERSONALITY NEEDS AND SOCIAL BACKGROUND AND MILITARY STATUS VARIABLES\*

*Department of Psychology, San Diego State College*

MAXINE MURPHY GUNDERSON<sup>1</sup>

### A. INTRODUCTION

Relationships between sociocultural background and personality variables have been explored in a number of studies. Hearn *et al.* (4), Moldovan (7), and Byers *et al.* (2) have found personality needs, as measured by the Edwards Personal Preference Schedule (EPPS), to be related to recalled behavior early in life or to certain parent-child relations. Mehlman *et al.* (6) found the EPPS Dominance, Affiliation, and Nurturance Scale scores to be related to social class. Knowledge of interpersonal relationships in adulthood would be enhanced by further exploration of the relationship between personal need patterns and social background, as well as adult occupational and social status. The present study considers relationships between expressed personality needs as measured by the EPPS and demographic social background and military status variables.

### B. METHOD

#### 1. Subjects

Subjects were 219 navy enlisted volunteers for Antarctic service. From among these volunteers each year, men are selected to participate in the United States Antarctic Research Program sponsored jointly by the National Science Foundation and the United States Navy. The mean age of the sample was 26.1 years and the mean length of service was 6.9 years. Of the total sample 55 per cent had completed high school and 18 per cent had some college experience.

#### 2. Procedure

Source of social background data was a 16-page personal history booklet which provided a variety of information on military status, childhood resi-

\* Received in the Editorial Office, Provincetown, Massachusetts, on November 29, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> The assistance of Frank A. Thompson in data processing is gratefully acknowledged.



dence, socioeconomic background, occupational and educational history, religious affiliation and frequency of worship, family relationships, preferences for social contact, physical symptoms, physical qualities, birth order, etc. The testing was done by the U.S. Navy Medical Neuropsychiatric Research Unit in San Diego. Subjects were tested as part of the routine screening for Operation Deep Freeze. Personal history and military status data were coded on IBM cards and intercorrelated with EPPS scale scores.

### C. RESULTS AND DISCUSSION

In this study each of the 15 EPPS variables was shown to have significant relationships with from one to 11 of the personal history variables, and 11 of the EPPS variables were significantly related to from three to 11 of the personal history and military status variables. Table 1 presents the correlation coefficients obtained for the EPPS scales and the 35 personal history and military status variables which yielded at least one significant relationship. The significant relationships will be presented and discussed for each of the 15 EPPS scales in order.

Achievement (ACH) need scale scores were significantly correlated with five of the personal history background variables: pay grade, college attendance, closeness to the mother, preference for being alone, and parental agreement. The positive correlation with pay grade supports the concurrent validity of the Achievement Scale: pay grade represents the rank or occupational status attained by the individual. The correlation between ACH and closeness to mother has relevance to Bradburn's (1) suggestion that boys with high achievement need had dominant mothers, while boys with low nAch had dominant fathers who tended to interfere with their son's attempts to achieve. The ACH relationship with the parental agreement variable suggests an association between achievement need and parental cohesiveness and unanimity.

Four background variables were found to be significantly related to the Deference (DEF) Scale. Age, years of service, and pay grade were all positively correlated with DEF. These correlations are consistent with attitudes of military conformity to be expected among experienced navy enlisted men. Father's occupation in a skilled trade was negatively correlated with DEF, suggesting that subjects who came from homes where the father was in a skilled trade apparently received a different orientation toward authority than those whose fathers were in other occupations.

The need for Order (ORD), or as Edwards (3) describes it, the need to stress organization and neatness in one's activities, was significantly correlated

with 11 of the personal background variables. Age and years of service were positively correlated with ORD. The association of years of service and orderliness would agree with the "spit and polish" stereotype of the navy. Being

TABLE 1  
CORRELATIONS<sup>a</sup> BETWEEN PERSONAL HISTORY VARIABLES AND EPPS SCORES

Variables	ACH	DEF	ORD	EXH	AUT
Pay Grade	17*	17*	04	—06	—06
Yrs. Serv.	—04	23***	18**	07	—22**
Age	04	24***	16*	—05	—22***
Educ. Coll.	13*	—05	11	01	04
Mobility	—08	—02	—00	—05	06
South	—00	—01	—01	—06	00
Size Comm.	—04	—10	—16*	05	—00
SES	03	—07	—22***	04	05
Married	04	—02	03	—01	—01
Div., Sep.	09	08	01	—03	—12
Fam. Size	06	09	07	—02	—05
Father Educ.	02	—04	—19**	—00	00
Oldest	—00	—06	—15*	02	10
Middle	00	04	05	04	—15*
Close M	16*	—09	—06	03	01
Close F	—05	02	—03	09	—05
Close Both	—06	—01	—01	—13	—05
Close Neither	06	08	14*	01	—04
P. Agree	14*	00	—11	04	—13
Fam. Squab.	02	—10	—22***	—06	02
Fa. Infl.	01	—02	05	14*	05
Catholic	—07	02	04	—03	01
Neither Prot. nor Cath.	01	—03	—15*	—02	03
P. Rel. Diff.	—06	—03	—01	—12	03
Worship	—12	—10	02	06	—09
Symptoms	00	—08	—14*	—13	06
Phys. Qual.	11	—05	—06	08	—12
Hooky	—07	—04	10	—12	04
Clubs	—03	06	—07	03	—02
Books	08	01	01	03	01
F. Clk., Sales	09	—04	—14*	05	—01
F. Skilled	01	—17*	—01	10	—02
Prefer to be with others	01	—04	—11	—04	—20**
Prefer to be alone	15*	—02	12	03	14*
Prefer neither	—08	05	06	03	14*

Note: EPPS = Edwards Personal Preference Schedule; ACH = Achievement; DEF = Deference; ORD = Order; EXH = Exhibition; AUT = Autonomy; AFF = Affiliation; INT = Intracception; SUC = Succorance; DOM = Dominance; ABA = Abasement; NUR = Nurturance; CHG = Change; END = Endurance; HET = Heterosexuality; and AGG = Aggression.

<sup>a</sup> Decimal points have been omitted.

\*  $p < .05$ .

\*\*  $p < .01$ .

\*\*\*  $p < .001$ .

TABLE 1 (continued)

Variables	AFF	INT	SUC	DOM	ABA
Pay Grade	—06	—10			
Yrs. Serv.	—09	—06	02	38***	—39***
Age	—06	—06	05	24***	—11
Educ. Coll.	06	02	08	28***	—17*
Mobility	—01	11	—08	18**	—16*
South	04	—00	—01	—18*	11
Size Comm.	—04	—05	—01	—01	—02
SES	13	04	12	20**	—05
Married	—03	—03	—01	10	04
Div., Sep.	02	14*	—00	16*	—11
Fam. Size	—06	—09	—06	02	—02
Father Educ.	15*	05	—06	01	00
Oldest	15*	06	—01	14*	—10
Middle	—08	—13	—02	05	—10
Close M	—06	—04	03	02	04
Close F	08	—10	00	07	—09
Close Both	—03	17*	01	—13	18**
Close Neither	02	07	—03	09	—00
P. Agree	10	03	—12	02	—10
Fam. Squab.	01	—00	01	00	05
Fa. Infl.	—01	—04	—08	03	—00
Catholic	—16*	01	—10	02	02
Neither Prot.			13	05	06
nor Cath.					
P. Rel. Diff.	01	—01	—02	—03	00
Worship	03	05	—14*	02	—04
Symptoms	04	03	05	06	15*
Phys. Qual.	05	03	—03	—13	03
Hooky	06	04	—10	14*	—14*
Clubs	—12	03	02	—01	15*
Books	01	06	—05	15*	—05
F. Clk., Sales	—01	—04	—06	13	—16*
F. Skilled	—02	03	07	11	03
Prefer to be	05	—04	—01	—04	01
with others					
Prefer to be	20**	10	16*	08	—05
alone	—14*	01	—16*	—04	—02
Prefer neither	—15*	10	—10	—06	06

close to neither parent was positively correlated with ORD. Socioeconomic (SES) status, father's education, being the eldest sibling, and presence of family friction and squabbles were all found to be negatively correlated with ORD. Two personal history variables, religious affiliation other than Catholic or Protestant, and father's occupation in clerical or sales, were negatively correlated with ORD.

Only one personal background variable was significantly associated with the Exhibition Scale, and this result can be attributed to chance.

The Autonomy (AUT) Scale, thought by Edwards to indicate a need to

TABLE 1 (continued)

Variables	NUR	CHG	END	HET	AGG
Pay Grade	—06	09	—02	04	—04
Yrs. Serv.	01	—11	10	—07	—04
Age	02	—08	06	—11	—07
Educ. Coll.	04	09	—07	—05	—04
Mobility	10	08	—02	—04	—03
South	14*	—10	—04	06	04
Size Comm.	04	00	—02	01	02
SES	00	14*	—06	—01	—11
Married	—01	00	10	—09	—01
Div., Sep.	—03	02	09	02	07
Fam. Size	00	—18**	04	05	12
Father Educ.	01	17*	—07	08	—07
Oldest	08	05	—16*	—01	—01
Middle	—03	—04	14*	02	07
Close M	—10	—07	—07	15*	08
Close F	—03	13	—06	07	10
Close Both	13*	—17*	01	—02	04
Close Neither	11	—01	14*	01	02
P. Agree	06	—08	02	02	02
Fam. Squab.	—05	05	—12	13*	01
Fa. Infl.	—11	14*	—06	01	—06
Catholic	01	—09	—03	—07	18**
Neither Prot. nor Cath.	09	12	—05	02	02
P. Rel. Diff.	—03	12	04	07	02
Worship	07	—13	07	—17*	06
Symptoms	—07	00	02	19**	10
Phys. Qual.	04	—02	—06	04	01
Hooky	10	—02	03	06	01
Clubs	—06	08	07	—10	—02
Books	06	06	02	—02	—11
F. Clk., Sales	—10	—03	00	06	00
F. Skilled	07	01	01	01	02
Prefer to be with others	06	—05	—06	14*	—06
Prefer to be alone	—08	—00	01	—07	09
Prefer neither	—03	05	06	—12	03

act independently of others and social conventions, was correlated significantly with six background characteristics. Age and years of military service were both negatively correlated with AUT, again reflecting a general association of advancing age and experience with conformity and acceptance of authority and interdependence. Being a middle child was negatively correlated with AUT. Preferring to be with others was also negatively correlated with this scale. Conversely, preferring to be alone or no preference was positively correlated with AUT.

The Affiliation (AFF) need was positively correlated with being the eldest child. This result conforms Schachter's (8) findings of greater affiliation need

for firstborn children. Other ordinal positions were not correlated with AFF. Consistent with expectations, the item "prefer to be with others" was positively correlated with AFF. Conversely, the "prefer to be alone" and no preference were negatively correlated with AFF. Catholic religious affiliation was negatively correlated with this variable.

Personal history relationships with Intraception, thought to measure the need to analyze one's own and other's motives, were positive for parents divorced or separated and closeness to both parents. There were three significant relationships between personal background characteristics and Succorance, or need to receive encouragement, sympathy, and affection from others. Preferring to be with others was positively correlated with this variable, while preferring to be alone was negatively correlated. The presence of religious differences between parents was also negatively correlated with Succorance.

Ten background characteristics were significantly correlated with the Dominance (DOM) need. Positive correlations with DOM were obtained for pay grade, years of service, and age. This pattern of results supports the validity of the DOM scale, since this variable would be associated with military rank and assertive and controlling behavior. Marriage, which is highly correlated with age, was positively correlated with the DOM scale. College attendance, father's education, subject's membership in clubs during school years, superiority of physical abilities, and size of community of residence were all positively correlated with DOM. Frequency of changes of residence (mobility) was negatively correlated with this variable.

Abasement (ABA) need, or the need to feel guilty when things go wrong, was found to be significantly correlated with eight of the social background variables. Age and pay grade were both negatively correlated with this variable, pay grade being the more discriminating. Attendance at college, reading many books, and the physical abilities index were also negatively correlated with ABA. Closeness to father was positively correlated with abasement, indicating that closer affectional ties with father were associated with submissiveness. Playing hooky from school was also positively associated with the Abasement score. Frequency of religious worship was also positively correlated with ABA.

The Nurturance scale was significantly correlated with only two variables, being close to both parents and being from the Southern United States.

The Need for Change Scale, or need to seek new experiences and to avoid routine, was positively correlated with father's education and family socioeconomic status; while family size was negatively correlated with need for



Change. Closeness to both parents, rather than closeness to either father or mother individually, was also negatively associated with the Change scale.

Birth order and indicated closeness to neither parent were correlated with the Endurance (END) scale, or need to work hard and keep at a task until completed. Being the oldest child was negatively correlated with END, while being the middle child was positively associated with the END scale score. An indication of closeness to neither parent was also positively associated with endurance.

Four variables were associated with Heterosexuality (HET) need. Positively correlated with this need were report of being closest to mother, incidence of family friction and squabbles, and preferring to be with others rather than alone. Report of physical symptoms was also correlated positively with HET. Report of frequency of religious worship was negatively correlated with this need.

Only one social background variable was significantly correlated with the Aggression scale and this could probably be attributed to chance.

#### D. SUMMARY

The correlations obtained in this study were generally low, but many significant and meaningful relationships appeared. Certain of the present findings have relevance to previous reports and to generally held hypotheses concerning family and social influence. Navy rank, or pay grade, was positively associated with ACH, DOM, and DEF; and negatively correlated with ABA. These results conform with characteristics of the military leader's role and provide evidence of concurrent validity for these EPPS scales. As might be expected, age and naval experience were positively correlated with Deference, Order, and Dominance needs; and negatively correlated with the need for Autonomy.

Being the oldest sibling was significantly related to the Affiliation need, which confirms Schacter's findings of affiliative behavior in firstborn children. Reported closeness to the mother was positively related to Achievement, while closeness to the father was significantly related to Abasement. These findings might have relevance to McClelland's (5) suggestion that the more sons felt loved and accepted by their fathers the lower was their achievement need.

In conclusion it appears that the results of this study of relationships between personal history data and the 15 expressed need scales of the Edwards Personal Preference Schedule provide some confirmatory evidence for the importance of certain childhood and social background experiences in the development of personality.

## REFERENCES

1. BRADBURN, N. M. n-Achievement and father dominance in Turkey. *J. Abn. & Soc. Psychol.*, 1963, **67**, 464-468.
2. BYERS, A. P., FORREST, G. G., & ZACCARIA, J. S. Recalled early parent-child relations, adult needs, and occupational choice. *J. Counsl. Psychol.*, 1968, **15**(4), 324-328.
3. EDWARDS, A. L. Edwards Personal Preference Schedule Manual. New York: Psychological Corp., 1959.
4. HEARN, J. L., CHARLES, D. C., & WOLINS, L. Life history antecedents of measured personality variables. *J. Genet. Psychol.*, 1965, **107**, 99-110.
5. MCCLELLAND, D. C., ATKINSON, J. W., CLARK, R. A., & LOWELL, E. L. The Achievement Motive. New York: Appleton, 1953.
6. MEHLMAN, M. R., & FLEMING, J. C. Social stratification and some personality variables. *J. Genet. Psychol.*, 1963, **69**(1), 3-10.
7. MOLDOVAN, S. Some familial antecedents of personality needs. *Bull. Maritime Psychol. Assoc.*, 1963, **12**(1), 28-40.
8. SCHACHTER, S. The Psychology of Affiliation. Stanford, Calif.: Stanford Univ. Press, 1959.

*Department of Psychology*  
*San Diego State College*  
*San Diego, California 92115*

## EVALUATION OF SELF AND OTHERS AND SOME PERSONALITY CORRELATES\*<sup>1</sup>

*Department of Psychology, Brigham Young University*

---

DARHL M. PEDERSEN

---

### A. INTRODUCTION

In recent years there has been a marked increase in the number of studies relating to the self-concept. A comprehensive review of the studies reported prior to 1961 has been given by Wylie (11). Much of the self-concept research has attempted to relate the phenomenal self to other variables. An examination of the empirical literature reveals a bewildering array of concepts measured, measuring instruments, hypotheses, and research designs used to accomplish this. The related concepts of self-esteem, self-acceptance, self-concept, self-evaluation, and discrepancy between self and ideal self have been measured by such techniques as Q-sorts, semantic differentials, and other rating scales, questionnaires, and adjective checklists. Also, the concepts have been related to such variables as body characteristics, success and failure experiences, participation in psychotherapy, acceptance of others, level of aspiration, parental characteristics, personality characteristics, etc. Because of the differences in subjects, measuring instruments, and procedures used from one study to another, the results have often been confusing and contradictory.

The present study attempts (a) to clarify previously studied relationships by examining several of them with the use of the same methodology with the same set of subjects and (b) to study a number of new relationships which are logically derived from those previously studied. A measuring instrument which meets the needs of these objectives is the semantic differential (8). With it the same rating scales may be used to obtain the subjects' ratings of a number of different concepts. It measures three dimensions of meaning—evaluation, activity, and potency. By the selection of scales which have large loadings on the evaluation dimension, a measure of evaluation or the relative "goodness"

---

\* Received in the Editorial Office, Provincetown, Massachusetts, on December 4, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This research was supported in part by a Faculty Research Fellowship granted by Brigham Young University. Appreciation is extended to Mr. Larry W. Bailey and Mrs. Sylvia B. Lassig for their assistance in the literature review and the data collection, respectively.

of a number of concepts may be obtained. In the present study evaluations of three types of concepts will be made: (a) evaluations of self (Self-Concept and Ideal Self), (b) evaluations of others (Generalized Other, Father, and Mother), and (c) perceived evaluations by others (Father's Perception of Me and Mother's Perception of Me). Evaluation differences scores will be obtained for pairs of these concept evaluations. And a number of personality variables will be measured. The purpose of the study is to determine the relationships that exist among these three types of variables.

## B. METHOD

### 1. Subjects

The subjects were 150 students (69 males and 81 females) enrolled in introductory psychology at Brigham Young University.

### 2. Measuring Instruments

*a. Self and Others Rating Scale (SORS).* The SORS was a semantic differential consisting of 25 bipolar adjectives which loaded highly on the "evaluative" (i.e., good-bad) dimension of meaning (8). They were separated by an eight-point scale. The 25 rating scales were used to describe each of the following concepts: Self-Concept (yourself as you really are—SC), Ideal Self (the kind of person you would like to be—IS), Generalized Other (other people in general—GO), Father (F), Mother (M), Father's Perception of Me (yourself as you think your father would rate you—FPM), and Mother's Perception of Me (yourself as you think your mother would rate you—MPM). Since each of the adjectives loaded highly on the evaluative dimension, and since there were no adjectives representing the dimensions of "activity" and "potency," the SORS measured primarily the evaluation of each concept. For example, the measure of self-concept represents evaluation of self-concept or self-evaluation.

The 25 bipolar adjectives in the order in which they appeared were as follows: rational-irrational, sacred-profane, graceful-awkward, moral-immoral, unintelligent-intelligent, awful-nice, worthless-valuable, unfair-fair, unselfish-selfish, unsuccessful-successful, unimportant-important, wise-foolish, unsociable-sociable, clean-dirty, beautiful-ugly, cruel-kind, bad-good, dishonest-honest, low-high, disreputable-reputable, wholesome-unwholesome, unpleasant-pleasant, grateful-ungrateful, optimistic-pessimistic, and sane-insane. By random assignment 12 of the favorable adjectives appeared at the left and 13 at the right. This was accomplished to discourage the subjects from using a response

set to mark a particular location on the scale. The eight points going from the unfavorable adjective to the favorable adjective of each pair were assigned the numerical values from one to eight. A score was obtained for each of the seven concepts rated on the SORS by summing the numerical values checked on all 25 scales. The possible score range for each concept was 25 to 200. Split-half reliability estimates were obtained for each of the seven scores by applying the Spearman-Brown formula to odd-even correlations. The reliability coefficients were as follows: Self-Concept (.92), Ideal Self (.85), Generalized Other (.96), Father (.94), Mother (.93), Father's Perception of Me (.93), and Mother's Perception of Me (.92).

*b. Biographical Data Sheet (BDS).* The subjects were asked to indicate their church activity at home (CAH) by marking a five-point rating scale and their age.

*c. Personality Inventory (PI).* The PI was constructed by Pedersen (9) on the basis of a study by Becker (2), which demonstrated that there was considerable equivalence of factors both within and between the personality inventories of Cattell and Guilford (3, 5, 6, 7). The PI has scales for the following six personality traits: Cycloid Disposition (C), measuring emotional instability; Rhathymia (R), measuring a happy-go-lucky attitude; Thinking Introversion (T), measuring an attitude of reflection and introspection; Cooperativeness (Co); Extroversion (E); and Neuroticism (N).

*d. California F Scale (A).* This test measured authoritarianism and consisted of 28 of the 30 items appearing in Forms 45 and 40 developed by Adorno, Frenkel-Brunswik, Levinson, and Sanford (1).

*e. Manifest Anxiety Scale (MAS).* The 50-item measure of manifest anxiety developed by Taylor (10) was administered.

*f. Adjective Check List (ACL).* Gough's (4) adjective checklist consisting of 300 adjectives was administered. Subjects checked those adjectives that were self-descriptive. Seventy-five of the adjectives have been classified as "most favorable." A score for Self-acceptance (SA) was obtained by counting the number of most favorable adjectives that the subject checked and dividing this by the total number of adjectives that were checked and expressing this quotient to the nearest percentage.

### 3. Procedure

The six measuring instruments were administered to the subjects during three 50-minute class periods, approximately one and a half weeks apart. Direct scoring yielded 18 scores for each person, including the seven obtained from the SORS. In addition a computer program was written which com-



puted 17 of the 21 possible algebraic difference scores among the seven SORS scores. An example of this type of score is the Ideal Self evaluation score minus the Self-Concept evaluation score (IS-SC). The four differences which were not obtained because they were not considered to be meaningful were FPM-GO, MPM-GO, M-FPM, and F-MPM. All 35 scores were then punched into IBM cards, and their means, standard deviations, and intercorrelations were obtained separately for each sex using an IBM 7040 computer.

### C. RESULTS AND DISCUSSION

#### 1. Means

The means, standard deviations, and significance of differences between means for males and females for each of the 35 variables are presented in Table 1. For each difference score the concept with the larger mean was arbitrarily placed on the left so that the means of all of the difference scores are positive.

TABLE 1  
MEANS, STANDARD DEVIATIONS, AND SIGNIFICANCE OF DIFFERENCES  
BETWEEN MEANS OF THE VARIABLES FOR MALES AND FEMALES

Variables	Males		Females		<i>t</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Self and Others Rating Scale					
1. Ideal Self—Concept (IS-SC)	24.86	15.00	26.54	11.86	— .77
2. Self Concept—Generalized Other (SC-GO)	26.09	19.66	20.40	19.00	1.80
3. Father—Self Concept (F-SC)	10.91	20.29	10.03	18.04	.28
4. Mother—Self Concept (M-SC)	16.86	19.96	17.54	12.68	— .25
5. Father's Perception of Me—Self Concept (FPM-SC)	7.61	14.39	8.08	12.66	— .22
6. Mother's Perception of Me—Self Concept (MPM-SC)	12.52	13.04	11.14	11.06	.70
7. Ideal Self—Generalized Other (IS-GO)	50.95	20.44	46.94	21.34	1.17
8. Ideal Self—Father (IS-F)	13.95	20.50	16.51	18.24	— .81
9. Ideal Self—Mother (IS-M)	8.00	16.67	9.00	10.47	— .45
10. Ideal Self—Father's Perception of Me (IS-FPM)	17.25	16.97	18.46	15.16	— .46
11. Ideal Self—Mother's Perception of Me (IS-MPM)	12.34	13.70	15.40	12.75	—1.42

Note: There were 69 males and 81 females. The *df* = 148 for each test.

\*  $p < .05$ .

\*\*  $p < .02$ .

\*\*\*  $p < .0005$ .

TABLE 1 (*continued*)

Variables	Males		Females		<i>t</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
12. Father—Generalized Other (F-GO)	37.00	17.64	30.43	25.07	1.83
13. Mother—Generalized Other (M-GO)	42.95	15.84	37.94	22.32	1.56
14. Mother—Father (M-F)	5.95	14.38	7.51	15.80	— .63
15. Father—Father's Perception of Me (F-FPM)	3.30	15.68	1.95	19.03	.47
16. Mother—Mother's Perception of Me (M-MPM)	4.34	13.83	6.40	13.94	— .91
17. Mother's Perception of Me—Father's Perception of Me (MPM-FPM)	4.91	14.19	3.06	10.76	.91
18. Self Concept (SC)	163.00	17.20	164.09	13.67	— .43
19. Ideal Self (IS)	187.86	12.60	190.63	11.11	—1.43
20. Generalized Other (GO)	136.91	17.53	143.69	20.19	—2.18*
21. Father (F)	173.91	18.14	174.12	21.47	— .06
22. Mother (M)	179.86	13.83	181.63	14.54	— .76
23. Father's Perception of Me (FPM)	170.61	17.68	172.17	17.75	— .54
24. Mother's Perception of Me (MPM)	175.52	13.24	175.23	15.14	.12
Biographical Data Sheet					
25. Age (AGE)	20.55	2.70	19.37	3.68	2.19
26. Church Activity at Home (CAH)	4.25	1.13	4.36	.95	— .65
Personality Inventory					
27. Cycloid Disposition (C)	15.87	8.69	18.63	7.91	—2.03*
28. Rhythymia (R)	30.54	8.70	29.33	9.82	.79
29. Thinking Introversion (T)	26.88	6.72	25.64	6.45	1.15
30. Cooperativeness (Co)	71.36	15.17	74.98	13.33	—1.55
31. Extroversion (E)	22.87	7.32	22.22	7.43	.54
32. Neuroticism (N)	14.93	8.60	18.48	8.26	—2.57**
Other scales					
33. California F Scale (A)	— 9.83	19.58	— 6.31	14.60	—1.26
34. Manifest Anxiety Scale (MAS)	10.13	7.22	14.67	8.16	—3.57***
The Adjective Check List					
35. Self-Acceptance (SA)	49.65	12.46	48.01	10.13	.89

The descending order of magnitude of the means of the seven concepts which was the same for both males and females was as follows: Ideal Self, Mother, Mother's Perception of Me, Father, Father's Perception of Me, Self-Concept, and Generalized Other. The evaluation of a particular parent was higher than the child's perception of that parent's evaluation of him. Evaluations involving mother were just under the Ideal Self and higher than evaluations involving father. Evaluation of the Self-Concept was the lowest ex-

cept for evaluation of the Generalized Other. Although females tended to evaluate the concepts higher than males, only the mean of the evaluations of the Generalized Other was significantly larger for females than for males ( $p < .05$ ). Females evaluated other people in general higher than males did.

The relative magnitudes of the means of the difference scores may be predicted from the order of magnitude of the single concepts. IS-GO had the largest mean of all the difference scores. The small means for FPM-SC and MPM-SC indicated that the level of the subject's self-evaluation tended to be close to his perception of the level of his parents' evaluations of him. Furthermore, evaluation of FPM was closer to SC than evaluation of MPM was. Evaluation of IS was about the same distance above SC as SC was above GO. The mother was evaluated closer to the ideal self than was the father for both sexes as indicated by the IS-F and IS-M means. The only mean difference score that was significantly different for males and females was that of MPM-GO. Males displayed a significantly larger discrepancy than females. This finding was primarily a result of the lower evaluation of GO by the males. Also it may be noted that MPM was the one concept that did not tend to be evaluated lower for males than for females. Only three differences between means for personality variables were statistically significant. Females had means which were significantly higher than those for males on the measures of neuroticism and anxiety—C, N, and MAS.

## 2. Correlations

Three different types of variables were included in this study. They were as follows: (a) evaluation difference variables (Nos. 1-17), (b) concept evaluation variables (Nos. 18-24), and (c) personality variables (Nos. 25-35). The following six types of relationships may be found among these three types of variables: (a) correlations of personality variables with concept evaluation variables, (b) correlations among concept evaluation variables, (c) correlations of personality variables with evaluation difference variables, (d) correlations of concept evaluation variables with evaluation difference variables, (e) correlations among evaluation difference variables, and (f) correlations among personality variables. Correlations among all of the variables have been given for males in Table 2 and for females in Table 3. The six types of relationships included in each table have been separated into six sections.

a. *Correlations of personality variables with concept evaluation variables.* Tables 2 and 3 may be examined to determine the personality variables that were significantly correlated with the evaluation of each of the seven concepts.

TABLE 2  
CORRELATIONS AMONG THE VARIABLES FOR MALES

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1.																	
2.	-33																
3.	36	-61															
4.	58	-68	75														
5.	33	-42	64	56													
6.	53	-45	56	72	47												
7.	42	72	-33	-23	-16	-04											
8.	38	36	-73	-32	-39	-17	63										
9.	21	52	-57	-68	-37	-39	65										
10.	60	07	-23	03	-55	07	50	72									
11.	59	07	-14	-06	-08	-37	50	66	50								
12.	04	41	47	10	27	14	43	-43	-08	59							
13.	32	38	18	42	19	35	60	05	-21	-19	-09						
14.	30	-08	-38	34	-12	22	14	59	-13	37	12	01	64				
15.	15	-41	71	45	-09	29	-28	-59	-40	21	-11	36	06	32			
16.	33	-56	55	76	37	10	-29	-30	-61	-02	27	01	27	-38			
17.	15	01	-13	09	-58	44	12	24	02	63	-26	-14	13	28	37		
18.	-70	55	-54	-73	-38	-65	02	02	25	-29	-15	-01	-24	-25	-35	-45	-20
19.	23	37	-32	-31	-13	-25	52	48	58	31	50	04	06	01	-29	-21	-10
20.	-32	-58	15	04	09	-13	-79	-39	-34	-36	-23	-47	-66	-15	11	19	-21
21.	-27	-16	60	14	35	01	-35	-79	-41	-53	-30	52	-02	-66	46	19	-34
22.	-04	-29	40	53	33	24	-31	-43	-67	-32	-27	14	31	17	21	54	-12
23.	-41	20	-01	-25	44	-25	-11	-29	-07	-74	-21	21	-07	-34	-41	-13	-67
24.	-39	28	-15	-24	-04	14	-02	-13	-07	-31	-56	13	04	-11	-16	-48	17
25.	-26	21	-27	-25	-04	-10	01	08	07	-19	-19	-08	-05	04	-31	-27	-05
26.	-14	-02	08	-02	29	-08	-12	-19	-10	-38	-08	08	-05	-15	-16	05	-37
27.	29	-12	14	30	-04	27	09	07	-10	28	05	03	23	22	22	18	29
28.	-07	-09	22	-02	12	-01	-14	-27	-04	-16	-07	15	-14	-34	18	-02	-13
29.	02	-09	-14	-02	-25	03	-07	15	04	23	-01	-26	-13	17	05	-06	28
30.	-24	12	-29	-39	-02	-33	-06	11	24	-20	05	-20	-34	-13	-35	-25	-28
31.	-32	07	04	-20	05	-16	-16	-27	-05	-32	-20	13	-15	-33	01	-13	-20
32.	22	-09	13	23	-10	31	07	03	-08	28	-06	05	18	13	26	04	38
33.	18	-11	10	22	12	10	03	04	-09	06	11	-01	13	16	01	22	-04
34.	44	-33	19	44	18	31	00	14	-14	23	18	-16	15	35	08	34	10
35.	-29	-06	04	-15	14	-10	-27	-25	-09	-38	-23	-02	-26	-26	-08	-12	-23

TABLE 2 (continued)

Variables <sup>a</sup>	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34
1.																	
2.																	
3.																	
4.																	
5.																	
6.																	
7.																	
8.																	
9.																	
10.																	
11.																	
12.																	
13.																	
14.																	
15.																	
16.																	
17.																	
18.																	
19.	53																
20.	36	11															
21.	34	15	51														
22.	19	21	51	63													
23.	66	41	43	62	45												
24.	66	44	34	45	48	61											
25.	22	-02	-03	-10	-10	17	18										
26.	33	28	34	41	38	56	35										
27.	-39	-19	-24	-21	-05	-41	-23	07									
28.	19	17	29	42	20	28	23	-19	-16								
29.	-06	-06	04	-22	-11	-26	-05	-01	50	-18							
30.	45	32	30	10	00	42	26	10	-02	39	-05						
31.	48	28	39	50	32	51	47	25	22	-48	04	-23					
32.	-28	-13	-18	-12	-02	-35	-06	08	57	-40	80	-09	29				
33.	-06	14	07	05	24	04	02	-19	-21	86	-16	45	-43	-34			
34.	-50	-17	-12	-27	01	-34	-35	-13	29	23	26	16	-42	12	31		
35.	30	07	37	33	17	41	30	-17	-05	63	-19	33	-49	-33	51	33	
								17	17	-27	13	-19	38	27	-19	-04	-37

Note: Decimal points are omitted. Correlations above .23 are significant at the .05 level, and correlations above .30 are significant at the .01 level.

<sup>a</sup> See Table 1 for description of variables.



TABLE 3  
CORRELATIONS AMONG THE VARIABLES FOR FEMALES

Variables	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1.																	
2.	-10																
3.	31	-08															
4.	64	-05	52														
5.	24	-09	27	10													
6.	38	-01	30	32	60												
7.	46	83	10	31	05	21											
8.	34	02	-79	-10	-11	-05	21										
9.	36	-06	-27	-49	14	05	15	51									
10.	58	00	02	41	-65	-20	32	36	16								
11.	60	-09	03	32	-30	-51	25	36	29	72							
12.	15	70	66	34	12	21	70	-55	-24	01	-05						
13.	28	82	22	53	-02	17	89	-04	-33	23	10	78					
14.	16	06	-73	21	-23	-09	14	82	-08	31	22	-48	17				
15.	14	-02	77	42	-41	-11	06	-67	-36	45	23	54	23	-54			
16.	28	-04	23	66	-38	-51	12	-05	-48	53	70	14	34	26	47		
17.	11	10	-01	21	-56	33	16	09	-12	56	-18	07	21	18	36	-07	
18.	-63	27	-10	-39	-09	-26	-11	-31	-24	-41	-36	13	01	-20	-04	-15	-16
19.	29	22	20	20	14	08	36	-01	09	11	20	32	30	-08	10	11	-08
20.	-33	-76	01	-22	03	-17	-86	-22	-11	-28	-16	-57	-77	-19	-01	-06	-21
21.	-14	10	77	18	17	08	01	-86	-38	-25	-20	63	19	-74	62	10	-11
22.	-04	21	35	50	00	03	17	-37	-65	-03	-06	41	47	00	33	44	03
23.	-32	14	11	-23	64	22	-05	-32	-08	-78	-49	19	-01	-31	-32	-38	-53
24.	-29	24	12	-12	35	49	05	-31	-18	-52	-70	27	13	-24	-12	-50	09
25.	02	01	-35	-07	-06	01	01	36	11	07	01	-25	-04	34	-29	-07	08
26.	-12	-03	18	-16	28	09	-09	-25	06	-32	-18	10	-12	-33	-02	-21	-24
27.	36	12	05	29	-02	-03	31	18	05	30	36	13	27	17	06	29	-01
28.	-25	10	-10	02	-08	-11	-05	-06	-31	-13	-14	00	09	13	-04	11	-02
29.	01	38	-11	00	-16	01	34	12	02	15	00	20	32	13	00	-01	20
30.	-11	-25	09	-14	07	01	-28	-08	-06	-15	-11	-13	-29	-21	03	-14	-07
31.	-33	12	-15	-18	-01	-12	-08	-06	15	-25	-20	-02	00	03	-14	-07	-12
32.	30	10	00	23	-02	-06	25	20	06	25	33	07	21	19	01	26	-04
33.	-13	-15	13	-03	-12	-33	-21	-22	-11	00	16	-02	-15	-18	21	23	-20
34.	30	-01	-04	22	04	09	16	23	08	20	20	-04	11	22	-06	13	04
35.	-38	-10	00	-16	-04	-20	-29	-24	-23	-26	-18	-07	-17	-13	02	01	-16

TABLE 3 (continued)

Variables <sup>a</sup>	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34
1.																	
2.																	
3.																	
4.																	
5.																	
6.																	
7.																	
8.																	
9.																	
10.																	
11.																	
12.																	
13.																	
14.																	
15.																	
16.																	
17.																	
18.																	
19.																	
20.																	
21.																	
22.																	
23.																	
24.																	
25.																	
26.																	
27.																	
28.																	
29.																	
30.																	
31.																	
32.																	
33.																	
34.																	
35.																	

Note: Decimal points are omitted. Correlations above .22 are significant at the .05 level, and correlations above .28 are significant at the .01 level.  
<sup>a</sup> See Table 1 for description of variables.

After consideration was given to each concept in turn, it was possible to characterize what the person was like who evaluated each of the seven concepts highly. For each concept a comparison will be made between males and females.

Males and females high in self-evaluation (SC) were likely to be high in self-acceptance (SA: .30, .38) and low in anxiety (MAS: —.50, —.31). In addition the male was likely to be more active in church (CAH: .33), less neurotic (C: —.39; N: —.28), more cooperative (Co: .45), and more extroverted (E: .48). Failure to find significant correlations between neuroticism (as measured by C and N) and evaluation of self-concept for females indicated that they could be relatively neurotic and still evaluate themselves highly.

Males high in evaluation of Ideal Self (IS) tended to be more active in church (CAH: .28), more cooperative (Co: .32), and more extroverted (E: .28). None of the personality variables was related for females. It is interesting that church activity was related to how "good" the Ideal Self was for males, but it was not related to the evaluation of Ideal Self for females. Also neuroticism was not related to IS-evaluation for either males or females; therefore, the neurotic person was likely to have a low Self-Concept but not necessarily a low Ideal Self. Furthermore, self-acceptance (SA) was unrelated to how "good" a person rated his Ideal Self.

Males and females who were high in their evaluation of other people in general (GO) tended to be emotionally stable (C: —.24, —.22), cooperative (Co: .30, .29), and self-acceptant (SA: .37, .35). Also, males tended to be active in church (CAH: .34), happy-go-lucky (R: .29), and extroverted (E: .39); whereas females also tended to be introspective (T: —.31) and authoritarian (A: .25).

Males and females who evaluated their father (F) highly tended to be more active in church (CAH: .41, .25), lower in manifest anxiety (MAS: —.27, —.23), and more self-acceptant (SA: .33, .24). In addition, males tended to be more happy-go-lucky (R: .42) and extroverted (E: .50); and females tended to be younger (AGE: —.29) and more authoritarian (A: .22).

More variables were related to evaluation of father than to evaluation of mother (M). It appears that males and females tended to evaluate their mothers highly irrespective of their own personality characteristics. Males who were high in evaluation of their mothers tended to be more active in church (CAH: .38), more extroverted (E: .32), and more authoritarian (A: .24). Females who were high in mother-evaluation tended to be more self-acceptant (SA: .22). It is interesting to note that the authoritarian females

evaluated their fathers highly and that the authoritarian males evaluated their mother highly.

A number of personality variables were related to both the ratings of father's perception of me (FPM) and the ratings of mother's perception of me (MPM) for males, but there were almost none for females. The males who perceived that their fathers and mothers evaluated them highly tended to be more active in church (CAH: .56, .35), less emotionally unstable (C: —.41, —.23; MAS: —.34, —.35), more happy-go-lucky (R: .28, .23), more co-operative (Co: .42, .26), more extroverted (E: .51, .47), and more self-acceptant (SA: .41, .30). Males with high FPM-evaluations also tended to be less meditative (T: —.26) and less neurotic (N: —.35). Females who thought that their fathers evaluated them highly tended to be more active in church (CAH: .33) and more self-acceptant (SA: .27). Females who perceived that their mothers evaluated them highly tended to have lower anxiety (MAS: —.22). Church activity at home was related to FPM-evaluation for both males and females. The Mormon culture is a patriarchal one in which it is important to be active in church. It may be that the child perceives that there is greater approval of him by his father when he conforms to his father's expectations regarding church activity. Self-acceptance was related to perceived evaluation by both parents for males and by father for females. Children who believe that their parents think well of them tend to think well of themselves.

*b. Correlations among concept evaluation variables.* Evaluation of the self-concept (SC) was related to evaluation of all of the other concepts for both males and females (with the exception of evaluation of mother for males). Both males and females with a high self-evaluation tended to have a high evaluation of their ideal self (IS: .53, .56), other people (GO: .36, .42), parents (F: .34, .55; M: —, .60),<sup>2</sup> and to think that his parents have a high evaluation of him (FPM: .66, .70; MPM: .66, .71). For males self-evaluation was related to evaluation of father but not of mother. It is likely that the males evaluated their mothers highly irrespective of their self-evaluation. For females there was a reverse tendency. Self-evaluation was more highly related to evaluation of mother than of father. The evaluation of self was not as closely related to the evaluation of parents for males as it was for females.

For both sexes a high ideal-self evaluation (IS) was not only related to high self-evaluation but also the perception of a high evaluation by parents (FPM: .41, .53; MPM: .44, .56). Females with a high IS-evaluation tended toward higher evaluation of parents (F: .53; M: .70); males did not. In particular,

<sup>2</sup> In those instances where pairs of correlations are given, if one of the correlations is not significant, a blank will be given to identify which of the two was not significant.

a female's IS-evaluation tended to be highly correlated with her evaluation of her mother. The female probably adopts her mother as her model for her own ideal self. No relationship was found for males between their evaluations of ideal self and father. Interestingly, a person's evaluation of what he would like to be (IS-evaluation) was unrelated to his evaluation of how good other people in general are (GO-evaluation).

People with high evaluations of other people (GO) not only tended to have high self-evaluations but also tended to have high evaluations of parents (F: .51, .28; M: .51, —)—with the exception of mother for females—and to perceive their parents as having high evaluations of them (FPM: .43, .34; MPM: .34, .26). Males tended to evaluate their parents more nearly the same as other people than females did. The lower correlations involving GO for females together with their higher means on GO indicated that perhaps they evaluated other people positively irrespective of their own personal characteristics.

For both sexes high correlations tended to exist among the four remaining concept evaluation variables—F, M, FPM, and MPM. Children tended to evaluate their parents about the same, and they tended to think that their parents evaluated them about the same. Furthermore, their evaluations of their parents tended to be highly related to how they perceived that their parents evaluated them.

*c. Correlations of personality variables with evaluation difference variables.* The discrepancy between the evaluations of ideal self and self-concept (IS-SC) has commanded more research attention than all of the remaining difference scores included in this study. IS-SC is often used to measure self-acceptance or adjustment—the lower the IS-SC discrepancy the higher the self-acceptance or adjustment. The well-adjusted male tended to be older (AGE: —.26), more emotionally stable (C: .29), more cooperative (Co: —.24), more extroverted (E: —.32), less anxious (MAS: .44), and more self-acceptant (SA: —.29). The well-adjusted female tended to be more emotionally stable (C: .36; N: .30), more happy-go-lucky (R: —.25), more extroverted (E: —.33), less anxious (MAS: .30), and more self-acceptant (SA: —.38).

SC-GO indicated the extent to which a person evaluated himself higher than other people. The male with a small discrepancy had higher anxiety (MAS: —.33); whereas, the female was less introverted (T: .38) and more cooperative (Co: —.25). Males did not evaluate the generalized other as favorably as females; therefore, when a male evaluates himself at about the same level as other people, it is possible that it produces anxiety.



Both males and females tended to evaluate their fathers more nearly the same as themselves (F-SC) as they got older (AGE: —.27, —.35). For males this resulted from a lowered evaluation of the father coupled with a raised evaluation of the self. For females it resulted primarily from a lowered evaluation of the father. In addition males with relatively higher F-SC scores tended to be less cooperative (Co: —.29).

There were more personality correlates with the discrepancy between evaluation of mother and self (M-SC) than with the discrepancy between evaluation of father and self. Both males and females with large M-SC scores tended to be emotionally unstable (C: .30, .29; N: .23, .23) and anxious (MAS: .44, .22). Additionally, the male with a larger M-SC score tended to be younger (AGE: —.25) and less cooperative (Co: —.39).

Males and females who perceived that their fathers evaluated them relatively more highly than they evaluated themselves (FPM-SC) tended to be more active in church (CAH: .29, .28). Males also tended to be less meditative and introspective (T: —.25). On the other hand, the male who perceived that his mother thought more of him than he thought of himself (MPM-SC) tended to be more emotionally unstable (C: .27; N: .31), less cooperative (Co: —.33), and more anxious (MAS: .31). And the female with a large MPM-SC discrepancy was likely to be less authoritarian (A: —.33).

The person with a large IS-GO score would like to be relatively better than other people are. This type of male was less self-acceptant (SA: —.27). This type of female was more emotionally unstable (C: .31; N: .25), more meditative (T: .34), less cooperative (Co: —.28), and less self-acceptant (SA: —.29).

Large IS-F and IS-M scores indicated that the person wanted to be relatively better than his father and mother, respectively. Males who evaluated their ideal selves as being relatively better than their evaluation of their fathers tended to be less happy-go-lucky (R: —.27), less extroverted (E: —.27), and less self-acceptant (SA: —.25). And females with relatively larger IS-F scores tended to be older (AGE: .36), less active in church (CAH: —.25), less authoritarian (A: —.22), more anxious (MAS: .23), and less self-acceptant (SA: —.24). The male with a relatively large IS-M score tended to be more cooperative (Co: .24). The female with a relatively large IS-M score tended to be less happy-go-lucky (R: —.31) and less self-acceptant (SA: —.23). The personality correlates with IS-F for males was similar to the personality correlates with IS-M for females.

The discrepancies, IS-FPM and IS-MPM, indicated the extent to which the individual wanted to be better than his perception of the evaluations of him by his father and mother. Males and females who had small IS-FPM discrepancy scores tended to be more active in church (CAH:  $-.38$ ,  $-.32$ ), more emotionally stable (C:  $.28$ ,  $.30$ ; N:  $.28$ ,  $.25$ ; MAS:  $.23$ ,  $-$ ), more extroverted (T:  $.23$ ,  $-$ ; E:  $-.32$ ,  $-.25$ ), and more self-acceptant (SA:  $-.38$ ,  $-.26$ ). There were fewer personality correlates of IS-MPM scores, and the pattern of correlations was quite different for males and females. Males with smaller IS-MPM scores tended to be more self-acceptant (SA:  $-.23$ ), and females tended to be more emotionally stable (C:  $.36$ ; N:  $.33$ ).

The size of the F-GO and M-GO scores indicated the extent to which an individual evaluated his father and mother more highly than people in general. The smaller a male's F-GO score, the more he tended to be reflective and introspective (T:  $-.26$ ), and the smaller a female's F-GO score, the older (AGE:  $-.25$ ) she tended to be. This was primarily a result of her decreased evaluation of her father with age. There were more personality correlates with M-GO than with F-GO for both males and females. For both sexes the person who evaluated his mother relatively closer to his evaluation of other people tended to be more emotionally stable (C:  $.23$ ,  $.27$ ) and more cooperative (Co:  $-.34$ ,  $-.29$ ). In addition males were more likely to have high self-acceptance (SA:  $-.26$ ) and females were less likely to be introspective (T:  $.32$ ).

Subjects tended to evaluate their mother higher than their father; however, in some instances the father was evaluated higher than the mother. The extent to which a particular subject evaluated his mother higher than his father was indicated by the M-F score. Males with high M-F scores tended to be less happy-go-lucky (R:  $-.34$ ), less extroverted (E:  $-.33$ ), more anxious (MAS:  $.35$ ), and less self-acceptant (SA:  $-.26$ ). Females with high M-F scores tended to be older (AGE:  $.34$ ), less active in church (CAH:  $-.33$ ), and more anxious (MAS:  $.22$ ).

The difference scores, F-FPM and M-MPM, indicated the extent to which the subject evaluated a parent higher than he perceived that the same parent evaluated him. For males the pattern of personality correlates was about the same for both F-FPM and M-MPM. Males with large F-FPM and M-MPM scores tended to be younger (AGE:  $-.31$ ,  $-.27$ ), less cooperative (Co:  $-.35$ ,  $-.25$ ), more neurotic (N:  $.26$ ,  $-$ ), and more anxious (MAS:  $-.34$ ). Females with large F-FPM scores tended to be younger (AGE:  $-.29$ ), and females with large M-MPM scores tended to be more neurotic

(C: .29; N: .26) and more authoritarian (A: .23). Notice that for females the size of the discrepancy between their evaluation of their father and their perception of their father's evaluation of them was unrelated to emotional instability and anxiety. It was not surprising that the authoritarian female evaluated an authority figure (i.e., her mother) much higher than her perception of that authority's perception of her. In fact the correlation between A and F-FPM for females and between A and MPM for males both fell just short of significance.

The MPM-FPM score represented the extent to which the subject perceived that his mother evaluated him higher than his father. Males with a large MPM-FPM score tended to be less active in church (CAH: —.37), more emotionally unstable (C: .29; N: .38), more meditative and introspective (T: .28), less cooperative (Co: —.28), and less acceptant of self (SA: —.23). For females the only variable that was related was church activity (CAH: —.24).

*d. Correlations of concept evaluation variables with evaluation difference variables.* This type of correlation was very similar to the type considered in the last section. However, evaluations of the seven concepts rather than the personality characteristics were related to the concept discrepancy scores. Correlations of a discrepancy between two concepts with either of the concepts involved in the discrepancy were not considered, since these were part-whole correlations and consequently were spuriously high.

The discrepancy between IS and SC described the extent to which an individual thinks he is what he ought to be. As indicated previously, it has commonly been used as a measure of self-acceptance. Males and females who were relatively more self-acceptant by this measure tended to evaluate other people highly (GO: —.32, —.33) and perceived their parents as evaluating them highly (FPM: —.41, —.32; MPM —.39, —.29). In addition, males with low IS-SC scores tended to evaluate their fathers more highly (F: —.27).

The SC-GO score represented the extent to which the subject evaluated himself more highly than he did other people. Both males and females with high SC-GO scores tended to have a high ideal self (IS: .37, .22) and tended to perceive that their mother had a high evaluation of them (MPM: .28, .24). In addition males tended to have a low evaluation of their mothers (M: —.29).

F-SC and M-SC indicated the extent to which a person evaluated his father and mother higher than himself. The male with a high F-SC score tended to have a low ideal self (IS: —.32) and a high evaluation of his mother (M: .40), and the female tended to have a high evaluation of her mother (M: .35).

The male with a high M-SC tended to have a low ideal self (IS:  $-.31$ ) and to perceive that both parents had a relatively lower evaluation of him (FPM:  $-.25$ ; MPM:  $-.24$ ), and the female tended to have a low evaluation of others (GO:  $-.22$ ) and to perceive that her father had a relatively low evaluation of her (FPM:  $-.23$ ).

Large FPM-SC and MPM-SC scores denoted that the person perceived that his father and mother evaluated him much higher than he evaluated himself. Males with a high FPM-SC score tended to evaluate both parents higher (F:  $.35$ ; M:  $.33$ ), and females with a high FPM-SC score tended to perceive that their mothers had a higher evaluation of them (MPM:  $.35$ ). Males with large MPM-SC scores tended to have lower ideal selves (IS:  $-.25$ ), higher evaluations of their mother (M:  $.24$ ), and tended to perceive that their fathers had lower evaluations of them (FPM:  $-.25$ ); females with large MPM-SC scores tended to perceive that their fathers had higher evaluations of them (FPM:  $.22$ ).

IS-GO indicated the extent to which an individual wanted to be better than his evaluation of how other people are. For males high IS-GO scores went with low evaluation of parents (F:  $-.35$ ; M:  $-.31$ ). For females IS-GO was not related to any of the other concept evaluations.

The higher the IS-F and IS-M scores, the more the individual wanted to be better than his father and mother. The higher the IS-F score, the lower the male evaluated other people (GO:  $-.39$ ), his mother (M:  $-.43$ ), and his father's perception of him (FPM:  $-.29$ ); and the lower the female evaluated her self (SC:  $-.31$ ), other people (GO:  $-.22$ ), her mother (M:  $-.37$ ), and her parents' perceptions of her (FPM:  $-.32$ ; MPM:  $-.31$ ). The higher the IS-M score, the higher the male evaluated his self (SC:  $.25$ ), and the lower he evaluated other people in general (GO:  $-.34$ ) and his father (F:  $-.41$ ); and the lower the female evaluated her self (SC:  $-.24$ ) and her father (F:  $-.38$ ).

IS-FPM and IS-MPM represented the extent to which the individual wanted to be better than his perception of his parents' evaluations of him. Both males and females with high IS-FPM scores tended to have low evaluations of themselves (SC:  $-.29$ ,  $-.41$ ), other people (GO:  $-.36$ ,  $-.28$ ), their fathers (F:  $-.53$ ,  $-.25$ ), and their mothers' perceptions of them (MPM:  $-.31$ ,  $-.52$ ). Males also had a low evaluation of their mothers (M:  $-.32$ ). The correlations were quite different for males and females with high IS-MPM scores. The higher the IS-MPM score, the lower the male evaluated other people in general (GO:  $-.23$ ) and his parents in particular (F:  $-.30$ ; M:  $-.27$ ). For the female, the higher the IS-MPM score, the



lower was her evaluation of her self (SC:  $-.36$ ) and her concept of her father's evaluation of her (FPM:  $-.49$ ).

The individual's concept of the extent to which his father and mother were better than other people in general was reflected by his F-GO and M-GO scores. F-GO did not correlate with any of the concepts for males, but females who viewed their fathers as much better than other people tended toward high evaluations of ideal self (IS:  $.32$ ), mother (M:  $.41$ ), and mother's perception of them (MPM:  $.27$ ). For males a high M-GO went with a low self-concept (SC:  $-.24$ ), and for females it went with a high ideal self (IS:  $.30$ ).

M-F indicated the degree to which an individual evaluated the mother higher than the father. The higher the male's M-F score, the lower he evaluated his self-concept (SC:  $-.25$ ) and his father's perception of him (FPM:  $-.34$ ). The higher the female's M-F score, the lower she evaluated both of her parents' evaluations of her (FPM:  $-.31$ ; MPM:  $-.24$ ).

The degree to which a parent was evaluated higher than he was perceived as evaluating the child in return was represented by F-FPM and M-MPM. For males a high F-FPM score tended to go with a low self-concept (SC:  $-.35$ ) and a low ideal self (IS:  $-.29$ ), and for females a high F-FPM score tended to go with a high evaluation of mother (M:  $.33$ ). Males with a high M-MPM score tended to have a lower self-concept (SC:  $-.45$ ), and females with a high M-MPM score tended to perceive that their fathers had low evaluations of them (FPM:  $-.38$ ).

Finally, males who thought that their mothers had a relatively more favorable perception of them than their fathers, MPM-FPM, tended to have a relatively lower evaluation of their fathers (F:  $-.34$ ).

*e. Remaining types of relationships.* The two remaining types of relationships were correlations among evaluation difference variables and correlations among personality variables. Neither of these will be discussed. The first type will not be discussed for a number of reasons. First, complex patterns of interdependency exist among difference variables which produce spuriously high correlations whenever a difference involving two variables is correlated with other differences involving one of the two variables. For this reason it was difficult to know which of the evaluation difference scores that correlated with a given difference score were meaningful. If all correlations among overlapping differences were eliminated, an interpretation using the remaining significant correlations likely would have been incomplete and distorted. Second, correlations among nonoverlapping, as well as overlapping, difference scores were attenuated. Difference scores that are notoriously unreliable themselves



were intercorrelated. Therefore, even the correlations among the nonoverlapping difference scores tended to be spuriously small. Finally, this type of relationship was not as meaningful as the other types discussed. The correlations have been presented in Tables 2 and 3 for the reader, and they may be examined with the above considerations in mind. Correlations among the personality variables will not be discussed, since they represented an incidental finding of the study. The reader may refer to these correlations to assist him in the interpretation of the correlations of the personality variables with the other variables.

#### D. SUMMARY

A measure of evaluation of self and others and some measures of personality were administered to 69 males and 81 females. The Self and Others Rating Scale (SORS) was a semantic differential which yielded an evaluation score for each of the following concepts: Self-Concept, Ideal Self, Generalized Other, Father, Mother, Father's Perception of Me, and Mother's Perception of Me. Also 17 evaluation difference scores were obtained from pairs of the seven SORS concept evaluation scores. The measures of personality were Personality Inventory (contained subscales for Cycloid Disposition, Rhathymia, Thinking Introversion, Cooperativeness, Extroversion, and Neuroticism), California F Scale, Taylor's Manifest Anxiety Scale, and Gough's Adjective Check List (scored for Self-acceptance). The *S*'s age and rating of his church activity were also obtained.

Analysis of the data yielded means, standard deviations, and correlations. A number of interesting and meaningful relationships were found within each of the following six types of correlations: (a) correlations of personality variables with concept evaluation variables, (b) correlations among concept evaluation variables, (c) correlations of personality variables with evaluation difference variables, (d) correlations of concept evaluation variables with evaluation difference variables, (e) correlations among evaluation difference variables, and (f) correlations among personality variables.

#### REFERENCES

1. ADORNO, T. W., FRENKEL-BRUNSWIK, E., LEVINSON, D., & SANFORD, N. *The Authoritarian Personality*. New York: Harper, 1950.
2. BECKER, W. C. A comparison of the factor structure and other properties of the 16 PF and the Guilford-Martin inventories. *Educ. & Psychol. Meas.*, 1961, 21, 393-404.
3. CATTELL, R. B., SAUNDERS, D. R., & STICE, G. *Handbook for the 16 Personality Factor Questionnaire*. Champaign, Ill.: Inst. Personal. & Ability Testing, 1957.

4. GOUGH, H. G. The adjective check list as a personality research technique. *Psychol. Rep.* (Monog. Suppl. 2), 1960, 6, 107-122.
5. GUILFORD, J. P. An Inventory of Factors STDCR. Beverly Hills, Calif.: Sheridan Supply, 1940.
6. GUILFORD, J. P., & MARTIN, H. G. The Guilford-Martin Inventory of Factors GAMIN: Manual of Directions and Norms. Beverly Hills, Calif.: Sheridan Supply, 1943.
7. ———. The Guilford-Martin Personnel Inventory: Manual of Directions and Norms. Beverly Hills, Calif.: Sheridan Supply, 1943.
8. OSGOOD, C. E., SUCI, G. J., & TANNENBAUM, P. H. The Measurement of Meaning. Urbana, Ill.: Univ. Illinois Press, 1957.
9. PEDERSEN, D. M. The measurement of individual differences in perceived personality trait relationships and their relation to certain determinants. Unpublished Doctoral dissertation, University of Illinois, Urbana, 1962.
10. TAYLOR, J. A. A personality scale of manifest anxiety. *J. Abn. & Soc. Psychol.*, 1953, 48, 285-290.
11. WYLIE, R. C. The Self Concept: A Critical Survey of Pertinent Research Literature. Lincoln, Neb.: Univ. Nebraska Press, 1961.

*Department of Psychology*  
*Brigham Young University*  
*Provo, Utah 84601*

## EFFECTS OF VICARIOUS VERBAL STIMULI ON CONDITIONING OF HOSTILE AND NEUTRAL VERBS\*

*Guilford County Mental Health Center and Guilford College*

---

HOYT MELVYN GILLEY

---

### A. INTRODUCTION

It has been found that modification of verbal behavior can occur through exposure to the verbal interaction of *E* and an experimental confederate (3, 7, 9). The *S* increases frequency of emission of the verbal response class for which the confederate receives praise from the *E*. Verbal conditioning, then, has been successfully accomplished by vicarious reinforcement.

A factor in verbal conditioning for which vicarious factors have not been investigated intensively has been the preconditioning situation. Solley and Long (11) noted that *Ss* who were engaged in conversation prior to the conditioning task conditioned better than did *Ss* who were not so treated. Kanfer and Karas (6) and Simkins (10) noted that not all preconditioning conversation facilitated verbal conditioning. These investigators administered fake performance tasks and either praised or criticized *Ss* for their performance. It was found that *Ss* who were criticized initially subsequently conditioned better than *Ss* who had been praised warmly. These findings were interpreted by Simkins (10) in terms of deprivation and satiation of social reinforcers. Verbal stimuli, such as "good" and "fine," were more effective rewards during the conditioning task to those *Ss* who had been deprived initially of praise than to those *Ss* who had been satiated with social reinforcers prior to the conditioning task. The study by Gewirtz and Baer (4) supports this hypothesis. They found that children who were satiated with social reinforcers prior to a conditioning task conditioned less well than children who had been deprived initially of social reinforcers.

The problem for the present study was to determine if preconditioning verbal stimuli, such as praise and criticism, would significantly affect verbal conditioning if they were administered vicariously prior to a verbal conditioning task, as well as during the conditioning task itself. Two verbal response

---

\* Received in the Editorial Office, Provincetown, Massachusetts, on December 11, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

classes were chosen, hostile and neutral verbs. This was because Simkins (10) found that *Ss* to whom approving verbal stimuli were presented prior to a conditioning task did not condition relative to a control group, while a comparable group in Kanfer and Karas' study (6) did condition to a slight extent. Since Kanfer and Karas used neutral learning material (pronouns), while Simkins used hostile verbs such as "stabbed" and "murdered," differences in response classes used might have accounted for the slight differences in findings. It was decided to employ female *Ss* in the present study, since Simkins (10) found that female *Ss* were much more sensitive to preconditioning verbal praise and criticism than were male *Ss*.

## B. METHOD

### 1. *Subjects*

Subjects were 90 female undergraduate students enrolled in introductory psychology courses at Florida State University. They received course credit for participation in the experiment.

### 2. *Experimental Design*

This study involved a modified Lindquist Type III factorial design (8, pp. 281-284) for three nonrepeated measures and one repeated measure. Prior to the experiment each *S* was assigned randomly to one of 18 possible treatment combinations in (a) three kinds of examiner behavior toward an experimental confederate during a preconditioning task, (b) three conditions of verbal stimuli during the verbal conditioning task, and (c) two critical response classes. There were five *Ss* in each combination, or a total of 90 *Ss*.

### 3. *Procedure*

When the *S* arrived at the experimental room she was asked to sit down at a desk placed approximately five feet from where another female "student" (an experimental confederate) was seated. The *E* sat behind a table facing them and said: "In this experiment both of you will be doing a sentence completion task. But first, you (indicating confederate) volunteered to take an additional test, an intelligence test we are standardizing at F.S.U. After this, both of you will perform on a sentence completion task which is related to intelligence." The fake testing of the confederate consisted of giving vocabulary, arithmetic, anagram, and cryptogram tests. During this procedure different verbal stimuli were presented to confederate, depending on the group to which *S* was assigned. Preconditioning verbal stimuli were similar to those of Simkins (10):

*a. Approving verbal stimuli.* The confederate was praised for her performance. After each section of the performance task the confederate was shown fake norms which showed that she scored as high as the top 25 per cent of F.S.U. students on that task. Confederate was told she was giving an "unusually fine performance," and was complimented on the quality of her vocabulary and speed of calculation.

*b. Disapproving verbal stimuli.* The confederate was criticized for her performance after each section of the task. She was told she was too slow and was giving an "unusually poor" performance. After each subtest confederate was told a numerical score and presented with fake norms which showed her score to be in the bottom 25 per cent of F.S.U. students.

*c. Neutral verbal stimuli.* No comment was made on confederate's performance during the fake performance task.

The fake testing lasted approximately 15 minutes. Following this, identical decks of 3" × 5" unruled index cards were presented to *S* and confederate. At the top of each card was a pronoun and below that two verbs in the past tense. One of the verbs was considered to be intensely hostile (e.g., "killed") and the other verb was neutral (e.g., "compared"). The list of 20 hostile and 20 neutral verbs was the same as used in the study by Buss and Durkee (2) from which Simkins (10) selected the hostile verbs used in his study. The hostile and neutral verbs were assigned randomly to different sides of the card to avoid a position effect. Each hostile verb and each neutral verb appeared once in each block of 20 trials. Each time the same hostile verb appeared it was presented with a different neutral verb. The instructions to *S* and confederate were a modification of those used by Simkins (10) and were as follows: "This is a test of intellectual flexibility. I am going to present to both of you a series of cards. At the top of each card is a pronoun and below that two verbs. You are to make up a sentence starting with the pronoun and using either one of the two verbs. The same verbs are likely to occur repeatedly, but the idea of this test is to see if you can make up different sentences even though you might have to use the same verbs. You will give alternate responses. You (indicating confederate) turn to your first sentence. This procedure will be followed for all of the cards. Any questions before we begin?"

Before beginning the task, each *S* was assigned to one of two conditions of critical response class. For one group of *Ss*, the hostile verb group, hostile verbs were the critical response class. For the other group, (neutral verb), critical responses were neutral verbs. The task proceeded according to one of three conditions of verbal stimuli given by *E*: 1) *Verbal stimuli/Confederate's responses*: The confederate emitted critical responses according to a prearranged



schedule so that 5, 8, 11, 14, and 17 randomly arranged critical responses were emitted in successive 20-trial blocks of responses given by confederate. A similar procedure was used by Marston and Kanfer (9), who increased critical responses of a confederate in a linear fashion over blocks of trials. In the present study each critical response of confederate was followed by *E*'s saying "good" or "that's fine." The *S* did not receive verbal stimuli for giving critical responses. Subject and confederate gave 100 responses each. 2) *Verbal stimuli/S's responses*: Confederates emitted critical responses on the same schedule, but did not receive verbal stimuli from *E* for giving critical responses. The *S* did not receive verbal stimuli during the first block of 20 trials, and the number of critical responses *S* emitted during this time constituted an operant level of responding. During the next 80 trials, *E* said "good" or "that's fine" each time *S* emitted a critical response. 3) *No verbal stimuli*: Confederates emitted critical responses on the same schedule, but neither confederate nor *S* received verbal stimuli from *E*.

### C. RESULTS

The results of the analysis of variance indicated that the main effects for conditioning verbal stimuli (stimuli given during the conditioning task), critical response class, and trials were significant beyond the .05 level ( $F$  values of 6.75, 123.66, and 5.31, respectively). Subjects emitted neutral verbs at a greater average overall frequency (59.8) than hostile verbs (40.2). Critical response class was seen to be involved in significant interactions with trials ( $F = 7.45$ ), with conditioning verbal stimuli ( $F = 3.13$ ), and in the second order interaction of these three variables ( $F = 2.71$ ). Figure 1 represents this latter relationship graphically.

It can be seen that when hostile verbs were critical, *S*s tended to increase responding over trials regardless of experimental condition. When neutral verbs were critical, there was a tendency to decrease emission of neutral verbs. This represents a corresponding increase in emission of hostile verbs. The exception to this was when the *S*s were told "good" or "fine" for emitting neutral verbs. These *S*s did show evidence of increasing the number of critical responses emitted over trials.

The variable preconditioning verbal stimuli was not significant as a main effect ( $F = .02$ ). It was also not significant in any single order interaction, and was significant only in a double order interaction with conditioning verbal stimuli and trials ( $F = 2.13$ ). Since these latter two variables also interacted significantly with response class, the effects of preconditioning verbal stimuli *per se* were judged to be quite weak and obscure in this experiment.

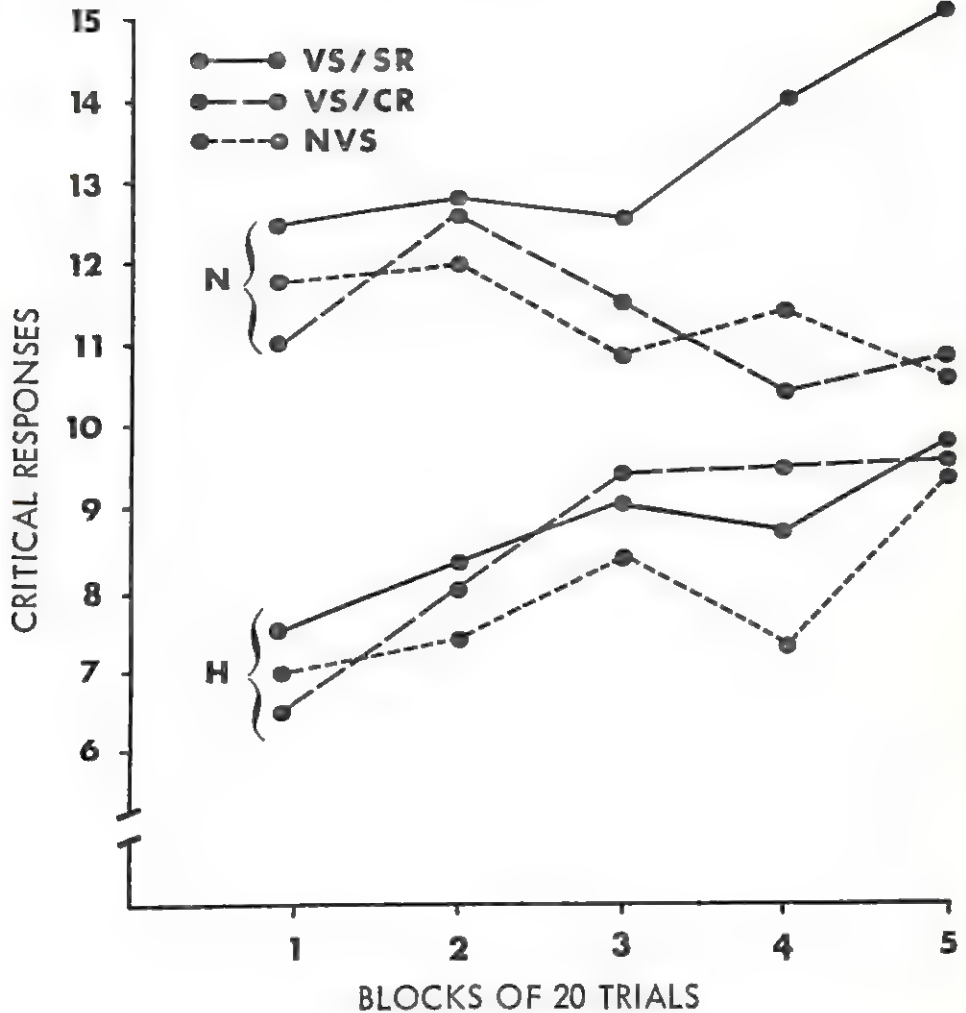


FIGURE 1  
FREQUENCY OF HOSTILE VERBS (H) AND NEUTRAL VERBS (N) OVER TRIALS UNDER CONDITIONS OF NO VERBAL STIMULI (NVS), VERBAL STIMULI CONTINGENT ON S'S RESPONSES (VS/SR), AND VERBAL STIMULI CONTINGENT ON CONFEDERATE'S RESPONSES (VS/CR)

#### D. DISCUSSION

From the results of previous studies on preconditioning verbal stimuli it had been predicted that disapproving verbal stimuli would facilitate verbal conditioning while approving verbal stimuli would retard conditioning (5, 10). The results of the present study did not support *this hypothesis*. Procedural differences might account for the discrepancy in results. In the studies

cited above the *Ss* were involved in a preconditioning task, while in the present study *Ss* watched another person perform the preconditioning task. Subjects in this study were noted to display a wide range of emotional expression while the *E* either praised the experimental confederate warmly or criticized her severely during the fake performance task. The varying degrees of amusement, apathy, and horror which different *Ss* displayed to identical stimuli would suggest that the vicarious effects of the preconditioning verbal stimuli in this study were not as consistent or as significant as the effects that would have been obtained had the *Ss* been engaged in the fake task and received verbal stimuli from the *E* directly.

The initial preference to emit neutral verbs rather than hostile verbs can be attributed most likely to the fact that college coeds were used as *Ss*. Buss and Durkee (2) found that female college students emitted neutral verbs with greater frequency than hostile verbs, both initially and over trials. Male *Ss* did not show a decided preference for either response class. Buss and Durkee explained these results in terms of differences in cultural training, which results in greater inhibition of hostile responses in women than in men.

The finding that *Ss* in the present experiment emitted an increasing number of hostile verbs over trials, except when told "good" and "fine" by the *E* for emitting neutral verbs, can possibly be explained on the basis of the characteristics and behavior of the experimental confederate. It seems likely that observing another female emit hostile verbs served to decrease the response inhibition which Buss and Durkee (2) suggested was associated with hostile responses in female *Ss*. Female *Ss* in Simkins' (10) experiment did condition when hostile verbs were critical. The baseline used for comparison was the performance of female *Ss* who received no verbal stimuli from *E*. Unlike the present procedure, however, *Ss* in Simkins' experiment were not exposed to the presence of another female who emitted hostile verbs. If Simkins had employed the present procedure it is possible that the control *Ss* in his study would have also shown an increase in the use of hostile verbs.

Subjects who were told "good" and "fine" by the *E* for emitting neutral verbs followed the traditional verbal conditioning paradigm in that they increased frequency of neutral verbs over trials. It seems likely that the direct verbal stimuli of the *E* was a more effective reinforcer for these *Ss* than was the use of hostile verbs by the confederate. In the other experimental conditions, the *S* had no direct feedback following emission of neutral verbs and had to rely on cues from the experimental confederate. Since the confederate emitted a variety of sentences with hostile content (with such hostile verbs as "mutilated," "strangled," "tortured," and "mangled"), the initial inhibition

of the *S* against using hostile verbs was lowered. This served to counteract the initial preference of the *S* for neutral verbs, whereas this initial preference was enhanced for those *Ss* who received verbal stimuli from the *E* for using neutral verbs.

It is recognized that *Ss'* initial preference for neutral verbs over hostile verbs and the change in response preference over trials could be interpreted in different ways. The characteristics of the *E*, for example, could be a critical factor. Binder, McConnell, and Sjöholm (1) found that a husky male *E* was unable to produce verbal conditioning in either male or female *Ss*, while an attractive female *E* successfully conditioned the *Ss*. The critical response class was hostile verbs. Greenspoon (5, p. 518) commented that the results of the study by Binder *et al.* (1) could be explained on the basis of a well-conditioned reaction in most people to inhibit hostile responses in the presence of husky males.

The significant interactions found in the present experiment indicated that the response measure used in verbal conditioning significantly affects the results and interpretation of the independent variables. It was concluded that the results of the present study should not be overgeneralized to male *Ss* or to other verbal response classes. It is obvious from the results of this experiment that the factors controlling verbal conditioning are complex and that the problem of identifying the actual stimulus in the conditioning situation is crucial.

#### E. SUMMARY

The present study was designed to investigate the effects of the following variables on verbal conditioning: (a) Observation of differential preconditioning verbal interaction of *E* and an experimental confederate. (b) Comparative conditioning effects of no verbal stimuli, verbal stimuli contingent on *S's* responses, and verbal stimuli contingent on responses of an experimental confederate. (c) Two critical verbal response classes. The *E* presented to a female confederate a fake performance task in *S's* presence prior to the conditioning task. The confederate was given either neutral, approving, or disapproving verbal stimuli during the preconditioning task. The verbal conditioning task consisted of having *S* and confederate alternate verbal responses on a task which consisted of choosing between a hostile and a neutral verb in composing sentences. Confederates emitted responses on a prearranged schedule. Experimenter said "good" or "fine" following critical responses of the *S*, the confederate, or gave no verbal stimuli, depending on experimental condition. No consistent effects of preconditioning verbal stimuli were noted. Neutral verbs were emitted at a significantly greater frequency than were hostile verbs.

All groups of Ss showed an increase in emission of hostile verbs, except for Ss who were directly reinforced for emission of neutral verbs. The results were seen to be determined in part by the initial inhibition of female Ss to emit hostile responses, and the subsequent modification of this tendency through imitation of the responses of the female experimental confederate.

## REFERENCES

1. BINDER, A., MCCONNELL, D., & SJOHOLM, N. A. Verbal conditioning as a function of experimenter characteristics. *J. Abn. & Soc. Psychol.*, 1957, **55**, 309-314.
2. BUSS, A., & DURKEE, A. Conditioning of hostile verbalizations in a situation resembling a clinical interview. *J. Consult. Psychol.*, 1958, **22**, 415-418.
3. FADIGAN, J. Indirect verbal conditioning. Unpublished M.A. thesis, Florida State University, Tallahassee, 1961.
4. GEWIRTZ, J. L., & BAER, D. M. Deprivation and satiation of social reinforcers as drive conditioners. *J. Abn. & Soc. Psychol.*, 1958, **57**, 165-172.
5. GREENSPOON, J. Verbal conditioning and clinical psychology. In A. J. Bachrach (Ed.), *Experimental Foundations of Clinical Psychology*. New York: Basic Books, 1962. Pp. 510-553.
6. KANFER, F., & KARAS, S. Prior experimenter-subject interaction and verbal conditioning. *Psychol. Rep.*, 1959, **5**, 345-353.
7. KANFER, F., & MARSTON, A. R. Human reinforcement: Vicarious and direct. *J. Exper. Psychol.*, 1963, **65**, 292-296.
8. LINDQUIST, E. F. Design and Analysis of Experiments in Psychology and Education. Boston, Mass.: Houghton Mifflin, 1953. Pp. 281-284.
9. MARSTON, A. R., & KANFER, F. H. Group size and number of vicarious reinforcements in verbal learning. *J. Exper. Psychol.*, 1963, **65**, 593-596.
10. SIMKINS, L. Effects of examiner attitudes and type of reinforcement on the conditioning of hostile verbs. *J. Personal.*, 1961, **29**, 380-395.
11. SOLLEY, C. M., & LONG, J. When is "uh-huh" reinforcing? *Percept. & Motor Skills*, 1958, **8**, 183-193.

3721 Manor Drive

Greensboro, North Carolina 27403



## THE EFFECTS OF ITEM DIFFICULTY SEQUENCING AND ANXIETY REACTION TYPE ON APTITUDE TEST PERFORMANCE\*<sup>1</sup>

*College of Education, The University of Oklahoma*

---

VINCENT F. BERGER,<sup>2</sup> DAVID C. MUNZ,  
ALBERT D. SMOUSE, AND  
HENRY ANGELINO

---

### A. INTRODUCTION

Traditionally, workers in the area of achievement and aptitude testing have recommended, without empirical justification, that the items comprising such tests be arranged in order of increasing difficulty: that is, the easier items first followed by the progressively more difficult ones (3, 5, 6, 7, 10, 11, 12, 15). On the basis of several arguments (8), test constructors have asserted that arranging test items in a sequence from easy-to-hard (E-H) eliminates adverse response sets which other item difficulty arrangements, such as hard-to-easy (H-E) or random (R), purportedly establish. Recent investigations of this standard test construction practice in connection with achievement testing have found no influence attributable to item difficulty sequencing when group measures are used (2, 14). However, Munz and Smouse (9), using a power achievement test in a classroom setting, found that individual reactions to the test-taking situation mediated the effect of item-difficulty sequencing on performance scores.

Although for aptitude testing there is some evidence to support the "superiority" (higher test scores) of the E-H arrangement of items over a subtest arrangement (13) and over an arrangement consisting of the hardest 25 per cent of the items followed by the remaining 75 per cent in an E-H order (8), no attempt has been made to investigate systematically the possible effects of E-H, H-E, and R item arrangements on aptitude test scores. This study was an attempt to evaluate the effect of such arrangements on aptitude test performance under both laboratory and field conditions. Further, in view of the

---

\* Received in the Editorial Office, Provincetown, Massachusetts, on December 12, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This study is based on a dissertation by the senior author submitted to the University of Oklahoma in partial fulfillment of the requirements for the Ph.D.

<sup>2</sup> Presently at the University of Pittsburgh, Pittsburgh, Pennsylvania.

recent research on achievement testing which suggests that achievement anxiety affects the relationship between item difficulty sequencing and achievement test performance (9), an attempt was made also to investigate the possibility that the relationship between item sequencing and aptitude test performance is similarly affected.

## B. METHOD

### 1. *Subjects*

Ss were 330 male and female high school seniors selected from three Oklahoma City schools. The three high schools had similar (a) general socioeconomic status and (b) proportions of college preparatory and noncollege preparatory students.

### 2. *Instruments*

Measures of the major independent variable, item difficulty sequencing, consisted of three arrangements of items taken from the Henmon-Nelson Test of Mental Ability, Revised Edition, grades 9-12.<sup>3</sup> The sequences were developed by reconstructing the original Henmon-Nelson Test, so that the 90 items, originally arranged in an order of increasing difficulty (E-H), were also arranged in the H-E and R orders. Thus, identical-appearing multilithed booklets contained the same multiple choice items and differed only as to item difficulty sequencing.

The second independent variable was the Achievement Anxiety Test (AAT), designed by Alpert and Haber (1) to measure the reported effects of anxiety experienced in test-taking situations. The AAT distinguishes between different degrees of anxiety reported by the respondent as either facilitating or debilitating to his test performance; each type of anxiety—i.e., facilitating and debilitating—is measured by a separate subtest of items. Since the authors of the AAT have preferred to view test-taking anxiety as a bidimensional construct which may have facilitating, as well as debilitating, effects on academic performance, four basic, achievement anxiety reaction types can be operationalized through use of the AAT. These types are (a) facilitators, those respondents scoring relatively high on the facilitating scale (AAT+) and relatively low on the debilitating scale (AAT—); (b) debilitators, those respondents scoring relatively high on AAT— and relatively low on AAT+; (c) nonaffecteds, those respondents scoring relatively low on both AAT+ and AAT—; and (d) high-affecteds, those respondents scoring relatively high on both AAT+ and AAT—.

<sup>3</sup> The authors wish to express their appreciation to the Houghton-Mifflin Company for their permission to use and alter the Henmon-Nelson Test of Mental Ability.

### 3. Procedure

Approximately two weeks prior to the actual experiment, the Ss were administered the AAT in the classroom by their regular teacher. In order to avoid the tendency of the students to give socially acceptable responses in answering the AAT items, while also taking into consideration the findings of Davids (+) that scores on anxiety scales are affected by the subjects' particular motivation for filling out anxiety questionnaires, the teachers were asked to read a set of instructions that emphasized that the questionnaire was "part of a research project" and "would not be used in any way to evaluate the students themselves."

Within each of the three schools, Ss were assigned to one of the two testing conditions (the "field" group and the "laboratory" group). In the "field" group the instructions and test setting were structured so as to parallel those conditions typically found in the actual aptitude-testing situation (e.g., the Ss were told that the test scores would become part of their permanent record and would be valuable in assessing future academic and occupational pursuits). In the "laboratory" group the instructions and test setting were structured so as to parallel those conditions found in the typical laboratory experiment (e.g., the Ss were told that they would be participating in a research project and therefore their test scores would not become part of their permanent record). Within each of the test conditions the three sequence forms were randomly distributed.

In order to test the major hypothesis that item difficulty sequencing influences aptitude test scores, the aptitude data (number of items answered correctly) for the 330 Ss were subjected to a 2 (test condition) by 3 (item sequence) analysis of variance.

In order to detect a possible interaction effect between achievement anxiety reaction and item difficulty sequencing on aptitude scores, the four anxiety reaction types based on the AAT were identified within each item-sequence by test-condition treatment combination following the typing procedure used by Munz and Smouse (9). This process eliminated 15 Ss from each of the six item-sequence by test-condition groups leaving data for a 2 (test condition) by 3 (item sequence) by 4 (achievement anxiety type) analysis of variance with a cell size of  $N = 10$  and a total  $N = 240$ .

### C. RESULTS

The first analysis, a 2 (test condition) by 3 (item sequence) analysis of variance of the aptitude data, revealed no statistically significant differences among the three item difficulty sequences ( $F < 1.00$ ,  $df = 2/324$ ,  $p > .05$ ).

The other main effect (test condition) also failed to produce a statistically significant  $F$  ratio ( $F < 1.00$ ,  $df = 1/324$ ,  $p > .05$ ). Moreover, the interaction of these two variables (test condition by item sequence) resulted in a non-significant  $F$  of 1.72 ( $df = 2/324$ ,  $p > .05$ ).

The second analysis, a 2 (test condition) by 3 (item sequence) by 4 (achievement anxiety type) analysis of variance, revealed no statistically significant differences among the three item sequences ( $F < 1.00$ ,  $df = 2/216$ ,  $p > .05$ ) or between the two test conditions ( $F < 1.00$ ,  $df = 1/216$ ,  $p > .05$ ). However, there was a significant effect of personality type on performance scores ( $F = 15.17$ ,  $df = 3/216$ ,  $p < .01$ ). Probing the significant effect of personality type on performance scores with the Newman-Keuls Test revealed that (a) facilitators scored significantly higher than debilitators ( $p < .01$ ), high-affecteds ( $p < .01$ ), and nonaffecteds ( $p < .05$ ); and (b) non-affecteds scored significantly higher than both debilitators ( $p < .01$ ) and high affecteds ( $p < .05$ ). None of the interactions was statistically significant ( $p > .05$ ).

#### D. DISCUSSION

In accordance with recent investigations in the achievement setting, the results of this study suggest that the standard test construction practice of arranging test items in an order of increasing difficulty is not justified in the aptitude setting, at least not from the argument that an E-H item difficulty sequence produces higher performance scores than an R or H-E arrangement. Moreover, while the results of this study do suggest differential individual reactions to the test-taking situation as measured by the AAT influence aptitude scores, it appears that this personality variable does not mediate the effect of item difficulty sequence on aptitude scores as has been shown with achievement scores. A possible explanation for the latter finding may be that the present study used a high school senior population instead of a college population as was used in the other related investigations. Since one might expect to find greater variation in intelligence, as well as test-taking ability and test-taking motivation in a high school population, it could be that this variability overshadowed any effects produced by an interaction between item sequence and achievement anxiety type. The possibility remains that such an interaction does exist and does affect aptitude test performance in college, or even college preparatory populations. Nevertheless, within the limits of this study, it appears that item difficulty sequence does not have an effect on high school aptitude scores.

In conclusion, recent research has yielded no justification for the test con-



structor's concern with the influence of item difficulty sequencing on group measures of intellectual aptitude and academic achievement for college and high school populations. However, there is evidence that personality variables interact with item sequencing in college achievement testing although this study did not find a similar effect on aptitude measurement in high schools.

### E. SUMMARY

To investigate the effects of item difficulty sequence and test condition (laboratory *versus* field) on aptitude scores, high school seniors were administered the Henmon-Nelson Test of Mental Ability reconstructed to consist of items ordered either hard-to-easy, easy-to-hard, or at random. Data were subjected to a 2 (test condition) by 3 (item difficulty sequence) analysis of variance which produced no significant  $F$  ratios. The effects of individual reactions to the test-taking situation were investigated by establishing four achievement anxiety reaction types based on the Anxiety Achievement Test. The data ( $N = 246$ ) were then subjected to a 2 (test condition) by 3 (item sequence) by 4 (reaction type) analysis of variance which yielded a significant  $F$  ratio ( $p < .01$ ) only for reaction type. The results of this study suggest that (a) item difficulty sequencing does not significantly affect aptitude test performance of high school seniors; and (b) differential reaction to test-taking anxiety does have a significant effect on aptitude test performance. Further implications of the research were discussed.

### REFERENCES

1. ALPERT, R., & HABER, R. N. Anxiety in academic achievement situations. *J. Abn. & Soc. Psychol.*, 1960, **61**, 207-215.
2. BRENNER, M. H. Test difficulty, reliability, and discrimination as functions of item difficulty order. *J. Appl. Psychol.*, 1964, **48**, 98-100.
3. CONRAD, H. S. The experimental tryout of test materials. In E. F. Lindquist (Ed.), *Educational Measurement*. Washington, D.C.: American Council on Education, 1951. Pp. 250-265.
4. DAVIDS, A. Relations among several objective measures of anxiety under different conditions of motivation. *J. Consult. Psychol.*, 1955, **19**, 275-279.
5. GREENE, H. A., & JORGENSEN, A. N. The Use and Interpretation of High School Tests. New York: Longmans, Green, 1939.
6. HELMSTADTER, G. C. Principles of Psychological Measurement. New York: Appleton-Century-Crofts, 1964.
7. JORDAN, A. M. Measurement in Education. New York: McGraw-Hill, 1953.
8. LUND, K. Test performance as related to order of item difficulty, anxiety, and intelligence. Unpublished Doctoral dissertation, Northwestern University, Evanston, Illinois, 1953.
9. MUNZ, D. C., & SMOUSE, A. D. Interaction effects of item difficulty sequence and achievement anxiety reaction on academic performance. *J. Educ. Psychol.*, 1968, **59**, 370-374.



10. NUNNALLY, J. C. *Tests and Measurement*. New York: McGraw-Hill, 1959.
11. REMMERS, H. H., & GAGE, N. L. *Educational Measurement and Evaluation*. New York: Harper, 1943.
12. ROSS, C. C. *Measurement in Today's Schools* (2nd ed.). New York: Prentice-Hall, 1947.
13. SAX, G., & CARR, A. An investigation of response sets on altered parallel forms. *Educ. & Psychol. Meas.*, 1962, **22**, 371-376.
14. SMOUSE, A. D., & MUNZ, D. C. The effects of anxiety and item difficulty sequence on achievement testing scores. *J. of Psychol.*, 1968, **68**, 181-184.
15. THORNDIKE, R. L., & HAGEN, E. *Measurement and Evaluation in Psychology and Education*. New York: Wiley, 1955.

*College of Education*

*820 South Oval*

*The University of Oklahoma*

*Norman, Oklahoma 73069*

## A METHODOLOGICAL NOTE ON "A SELF-CONCEPT INVENTORY"\*

*Department of Psychology, University of Manitoba*

---

SHIRIN SCHLUDERMANN AND EDUARD SCHLUDERMANN

---

Worchel (2) has developed a self-concept inventory for predicting maladjustment among cadets in military situations. Three sets of rating scores can be obtained; self, ideal, and other. Discrepancies between self-ideal, self-other, and ideal-other are taken as measures of adjustment or maladjustment. "Negative" ratings of self, ideal, and other are indicative of maladjustment. High discrepancies between self and ideal, as well as self and others, are indices of degrees of maladjustment. The normative sample consisted of groups of undergraduate and graduate students in psychology and education, as well as a group of 97 Air Force cadets. Further test-retest reliabilities of the inventory were studied in a sample of college students, which indicated that the self, ideal and other person rating scores are reliable measures. The inventory has been in use in research by investigators: e.g., Becker and Dileo (1).

The composite scores obtained from self, ideal, and other ratings are the basic units of analyses in derivation of discrepancy scores from the inventory. Somehow the implicit assumption is involved that ratings on self, ideal, and other are unitary and would yield meaningful composite scores. If the assumption is correct, then one would expect a few basic dimensions which would account for the maximum variance in three sets of ratings. In order to test this assumption, the authors decided to investigate the common factor space underlying self and ideal ratings through factor analysis. Worchel's self-activity inventory was administered with standard instructions to 177 college males and 150 college females enrolled in an introductory psychology course. The means and standard deviations of the replicated male sample were closely similar (see Table 1); in addition information was also obtained on the female sample which was not included in Worchel's normative sample.

Factor analysis was done by the Principal Component method, and rotated orthogonally by the Varimax solution. Four sets of factor matrices were obtained from four groups of  $54 \times 54$  correlation matrices: i.e., boys-self, boys-

---

\* Received in the Editorial Office, Provincetown, Massachusetts, on December 17, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

TABLE 1  
MEAN AND STANDARD DEVIATIONS OF THE COLLEGE SAMPLES

Measures	Replicated sample				Worchel's college sample	
	Girls		Boys		Boys	
	Self N = 150	Ideal N = 150	Self N = 177	Ideal N = 177	Self N = 153	Ideal N = 153
Means	139.28	97.50	140.16	100.90	136.5	103.9
Standard deviations	24.60	14.52	15.03	17.97	17.7	15.2

*Note:* According to the scoring system in the inventory, high scores on the self and ideal ratings are indicative of negative self and ideal images.

ideal, girls-self, and girls-ideal. Factor matrices consisted of 16 to 19 factors (in each matrix), which accounted for approximately 65 per cent to 68 per cent of the variance. If the proportion of variance accounted for by a factor or factors may indicate the difference between common and specific factors, then the evidence is that there are few common factors. The maximum variance accounted for by any one factor is 11 to 19 per cent. Because of the presence of a large number of specific factors, the assumption that self and ideal ratings are unitary may be questioned.

#### REFERENCES

1. BECKER, G., & DILEO, D. T. Scores on Rokeach's Dogmatism Scale and the response set to present a positive social and personal image. *J. Soc. Psychol.*, 1967, **71**, 287-293.
2. WORCHEL, P. Adaptability-screening of flying personnel: Development of a self-concept inventory for predicting maladjustment. School of Aviation Medicine, Report No. 56-62, U.S.A.F., Randolph Air Force Base, Texas, 1957.

*Department of Psychology  
The University of Manitoba  
Winnipeg, Canada*

# FACTORIAL STUDY AND ANALYSIS OF FACTOR VARIANCE OF INTELLECTIVE SELF-RATINGS FROM A COMBINED SAMPLE OF MALE AND FEMALE SUBJECTS\*<sup>1</sup>

*Stevens Institute of Technology*

---

RIO SCIORTINO

---

## A. INTRODUCTION

In three recent studies (7, 8, 9) the factor analyses of two Adjective Lists—the Motivation Adjective List (MAL) and the General Adaptability Adjective List (GAAL)—were reported. The MAL and the GAAL are two of a projected three adjective lists, constructed by Sciortino (6), to cover three areas of the intellectually superior and creative personality: namely, motivation, general adaptability, and intellection. A factor analysis (7) of the scores of a combined sample of male and female undergraduate Ss on the MAL yielded two factors: striving and assertion. An analysis of factor variance (8) on the two factors of MAL showed that the female Ss rated themselves significantly higher than the male Ss on both factors: striving and assertion. Another factor analysis (9) of the scores of a combined sample of male and female undergraduate Ss on the GAAL yielded five factors: sociability, energy, attentiveness, quickness, and articulateness. When, in the same study, an analysis of factor variance was performed on the five factors of GAAL, the results showed that the female Ss rated themselves significantly higher than the male Ss on factors sociability, energy, attentiveness, and quickness; no significant sex difference was found for factor articulateness.

The purpose of this study was: (a) first, to find the factorial structure of the Intellectual Adjective List (IAL) for a combined sample of male and female Ss; and (b) second, to determine, by analysis of factor variance, the effects of sex, male *vs.* female Ss, on the factor score means of the factors of the IAL obtained in (a). For the factorial structure of the IAL several factorial methods were used and the one that most adequately met the criteria of simple structure was chosen for the interpretation of factors. No correlations be-

---

\* Received in the Editorial Office, Provincetown, Massachusetts, on December 17, 1969, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> Data for this study were collected in the spring quarter of 1962 at the Ohio State University.

tween the scores on the IAL and scores on other instruments were undertaken in this study. In the analysis of factor variance of the IAL factors,  $F$  ratios were computed for the factor score means obtained from several factorial methods.

## B. FACTOR ANALYSES

### 1. Method

a. *Sample.* Male ( $n = 100$ ) and female ( $n = 102$ ) undergraduate students were required to serve as  $S$ s to fulfill the experimental requirements for a course in general psychology.

b. *Description of the test: The Intellectual Adjective List (IAL).* The IAL, constructed by Sciortino (6), consists of the 40 items listed in Table 1.  $S$ s rated themselves on each of the 40 items using the following five-point scale: (a) least, (b) less than moderately, (c) moderately, (d) more than moderately, (e) most. The self-ratings constituted the scores of  $S$ s on each of the 40 items. All the adjectives were defined and collected in a separate booklet entitled Adjective Definitions.

c. *Procedure.* The IAL and the Adjective Definitions booklet were given to  $S$ s at the same time.  $S$ s were asked to consult the Adjective Definitions booklet when they were unsure about the meaning of certain adjectives. The same examiner administered the IAL to both male and female groups on two different occasions. The findings reported here are based on the combined groups ( $n = 202$ ).

### 2. Results

a. *Computations.*<sup>2, 3</sup> The computations for this study were carried out on the IBM 7094 operated by the Princeton Computer Center and with the P-Stat computer program.

The iterative factor analysis routine of the P-Stat computer program was used to perform a factor analysis on the matrix of item correlations. The routine cycles on the number of significant roots and then iterates on the stability

<sup>2</sup> The following materials have been deposited with National Auxiliary Publication Service: Table A, matrix of item correlations ( $40 \times 40$ ) of the IAL; Table B, matrix of initial significant factors, and of quartimax, varimax and equamax factors; Table C, matrix of promax-quartimax and promax-varimax factors; and Table D, matrix of correlations of promax-quartimax and promax-varimax factors. Order NAPS Document No. 00231 from ASIS National Auxiliary Publications Service, c/o CCM Information Sciences, Inc., 22 West 34th Street, New York, N.Y. 10001; remitting in advance \$3.00 for photocopies or \$1.00 for microfiche.

<sup>3</sup> The data matrices of the male ( $n_1 = 100$ ) and female ( $n_2 = 102$ )  $S$ s on the 40 items (M) of the IAL were combined into an  $n_1 + n_2$  by M table to obtain the  $M \times M$  correlation matrix, the mean scores, and the standard deviations.



TABLE 1  
MATRIX OF MEANS, STANDARD DEVIATIONS, AND PROMAX-EQUAMAX FACTORS FOR SCORES  
ON THE INTELLECTIVE ADJECTIVE LIST (IAL) BY MALE ( $n = 100$ )  
AND FEMALE ( $n = 102$ ) Ss

Variables	M	SD	Promax-Equamax factors				
			A	B	C	D	E
1. Questing	3.26	.80	-.03	.06	.40	.12	-.08
2. Observant	3.72	.85	.06	.01	.04	-.11	.32
3. Interests-wide	3.60	1.01	.00	.07	.09	.05	.30
4. Explorative	3.32	.87	.13	.10	.30	.00	-.03
5. Curious	3.76	.81	.01	-.09	.67	.02	.01
6. Questioning	3.43	.88	.01	.01	.57	.02	.01
7. Open-minded	3.86	.79	-.12	.28	-.07	.09	.33
8. Likes challenges	3.62	.82	-.14	-.03	.27	.12	.25
9. Inquiring	3.42	.83	.01	.07	.61	-.02	-.03
10. Eager	3.60	.75	-.06	-.01	.05	.46	.13
11. Intelligent	3.26	.57	.05	.30	-.08	.28	-.01
12. Foresightful	3.20	.93	-.01	-.04	-.05	.08	.48
13. Clear-thinking	3.22	.70	-.02	.05	.08	.27	.19
14. Capable	3.60	.69	-.02	-.03	.05	.36	.14
15. Alert	3.50	.63	.12	-.21	.00	.49	.20
16. Keen	3.18	.64	.27	-.01	.00	.20	.10
17. Insightful	3.28	.75	-.02	.36	-.08	.02	.36
18. Flexible	3.59	.78	-.08	.04	.03	.04	.47
19. Clever	3.08	.70	.49	-.10	-.05	-.02	.24
20. Astute	2.84	.84	.11	.12	-.09	.19	.08
21. Productive	3.19	.71	.11	-.10	.12	.30	-.11
22. Inventive	2.68	1.07	.52	.01	.12	-.08	-.06
23. Ingenious	2.54	.91	.64	-.08	.03	-.04	-.06
24. Imaginative	3.34	.97	.32	.05	.11	-.12	.14
25. Adroit	2.92	.70	.56	-.13	-.09	.19	-.13
26. Resourceful	3.25	.70	.22	-.01	.12	.16	.19
27. Original	3.08	.83	.35	.15	-.10	-.01	.10
28. Innovative	2.99	.71	.38	.13	.01	-.01	-.01
29. Improvisive	3.29	.74	.32	.03	.05	.06	.10
30. Fertile (mind)	3.18	.77	.13	.29	-.08	.14	.15
31. Studious	3.08	.86	-.18	.08	.00	.46	-.07
32. Reflective	3.40	.90	-.02	.25	.25	-.09	.08
33. Pensive	3.11	1.03	-.02	.42	.19	-.03	-.08
34. Ideational	3.14	1.04	.12	.33	.14	-.09	.05
35. Academic	2.84	.96	.02	.42	-.07	.33	-.26
36. Theoretical	2.83	.99	.10	.44	.06	-.03	-.07
37. Scholarly	2.90	.89	-.05	.25	-.03	.38	-.01
38. Meditative	3.33	.89	-.21	.60	.01	-.04	.31
39. Introspective	3.37	.90	-.05	.43	-.06	-.06	.27
40. Esthetic	3.43	1.00	.13	.23	-.16	.02	.25
.10 h.c.			.18	.22	.26	.22	.19

Note: Initials h.c. stand for hyperplane count. Decimal points omitted.

of diagonal estimates. The obtained initial matrix of significant factors corresponds to the significant roots. Table A shows the matrix of item correlations.

Three orthogonal rotations of the initial matrix of significant factors were

performed using the quartimax (4), the varimax (2), and the equamax (5) procedures. Table B shows the matrix of the initial significant factors, and of the quartimax, varimax, and equamax factors. Three promax rotations (1), with  $k = 4$ , were performed on each of the three orthogonal rotations obtaining the promax-quartimax, the promax-varimax, and the promax-equamax solutions. Table C shows the matrix of promax-quartimax and promax-varimax factors, and Table D shows their corresponding correlations.

Table 1 shows the means and standard deviations of the IAL items, and the promax-equamax factors for the combined sample of 100 males and 102 females. Table 2 shows the correlations of the promax-equamax factors. Tables B and C, of the NAPS document, and Table 1, of the present study show that the total .10 hyperplane count is higher in the promax solutions than in the orthogonal solutions. Since the hyperplane counts of the promax-varimax and the promax-equamax solutions are numerically equal, 107 each, either one could be used to interpret the factors.

*b. Interpretation of factors.* Interpretation of factors is based on the results obtained from the promax-equamax solution, shown in Table 1, and rests on loadings of .25 or greater ( $p = .01$ , if  $n = 202$ ). A factor, whenever possible, was assigned the name of the variable that has the highest loading with it:

Factor A (Ingenuity) has loadings arranged as follows: ingenious, .64; adroit, .56; inventive, .52; clever, .49; innovative, .38; original, .35; imaginative, .32; improvisive, .32; and keen, .27. This factor refers to Ss' ingenuity or ability to solve problems in a skillful and inventive manner.

Factor B (Meditativity) has loadings arranged as follows: meditative, .60; theoretical, .44; introspective, .43; pensive, .42; academic, .42; insightful, .36; ideational, .33; intelligent, .30; fertile-mind, .29; open-minded, .28; reflective,

TABLE 2  
CORRELATION MATRIX OF PROMAX-EQUAMAX FACTORS FOR SCORES  
ON THE INTELLECTIVE ADJECTIVE LIST (IAL)  
( $n = 202$ )

Factors	A	Promax-Equamax factors			
		B	C	D	E
Ingenuity (A)		60*	47*	49*	
Meditativity (B)			50*	45*	44*
Curiosity (C)				44*	31*
Alertness (D)					36*
Adaptability (E)					28*

Note: Decimal points omitted.

\*  $p = .01$ .

.25; and scholarly, .25. This factor refers to Ss' meditativity or ability to reflect on the inner world of ideas.

Factor C (Curiosity) has loadings arranged as follows: curious, .67; inquiring, .61; questioning, .57; questing, .40; explorative, .30; likes-challenges, .27; and reflective, .25. This factor refers to Ss' curiosity or readiness to find out about things, in general.

Factor D (Alertness) has loadings arranged as follows: alert, .49; eager, .46; studious, .46; scholarly, .38; capable, .36; academic, .33; productive, .30; intelligent, .28; and clear-thinking, .27. This factor refers to Ss' alertness or ability to quickly size up situations, in general.

Factor E (Flexibility) has loadings arranged as follows: foresightful, .48; flexible, .47; insightful, .36; open-minded, .33; observant, .32; meditative, .31; interest-wide, .30; introspective, .27; academic, —.26; likes-challenges, .25; and esthetic, .25. This factor refers to Ss' flexibility or adjustment to environmental conditions, in general. The low negative loading of the item academic indicates that Ss' flexibility is of a behavioral rather than of a scholarly type.

## C. ANALYSIS OF FACTOR VARIANCE

### 1. Method

Computations:<sup>4</sup> The orthogonal and oblique factors of the IAL, reported above, were used in this section to obtain their respective factor scores. For this purpose the factor score subroutine of the P-Stat computer program, based on the multiple regression method, was used. The obtained factor scores were then subjected to a one-way classification analysis of variance with the use of another subroutine of the P-Stat computer program. Table 3 shows the means, standard deviations, and *F* ratios of the male and female Ss on the factor scores of the five factors of IAL (ingenuity, meditativity, curiosity, alertness, and flexibility), and for six factorial procedures (quartimax, varimax, equamax, promax-quartimax, promax-varimax, and promax-equamax).

### 2. Results

a. *Factor A (Ingenuity)*. In the quartimax, and promax-quartimax solutions the means of the female Ss are higher than those of the male Ss. In the varimax, equamax, promax-varimax, and promax-equamax solutions the means of the male Ss are higher than those of the female Ss. However, the *F* ratios reached statistical significance at the .05 level only in the varimax and the

<sup>4</sup>  $S = ZR^{-1}F$ ; where *S* = factor score matrix, *Z* = matrix of observed scores in standard score form, *R* = intercorrelation matrix, and *F* = rotated factor loading matrix.

TABLE 3

MATRIX OF MEANS, STANDARD DEVIATIONS AND *F* RATIOS FOR SCORES ON THE FIVE FACTORS OF THE IAL—INGENUITY, MEDITATIVITY, CURIOSITY, ALERTNESS, FLEXIBILITY—AND FOR SIX FACTORIAL MATRICES—QUARTIMAX, VARIMAX, EQUAMAX, PROMAX-QUARTIMAX, PROMAX-VARIMAX, PROMAX-EQUAMAX (Female  $n = 102$  and Male  $n = 100$ )

Factorial procedures	A Ingenuity		B Meditativity		C Curiosity		D Alertness		E Flexibility	
	F	M	F	M	F	M	F	M	F	M
<b>Quartimax</b>										
Means	5.04	4.96	5.03	4.97	5.14	4.86	5.12	4.88	5.40	4.59
SD	1.08	1.00	1.18	1.12	1.21	1.10	1.27	1.15	1.38	.89
<i>F</i> s		.33		.14		3.06		2.06		24.26***
<b>Varimax</b>										
Means	4.84	5.17	4.99	5.01	5.03	4.97	5.10	4.90	5.40	4.59
SD	1.11	1.10	1.20	1.09	1.18	1.09	1.29	1.09	1.34	.89
<i>F</i> s		4.49*		.02		.11		1.44		25.16***
<b>Equamax</b>										
Means	4.82	5.18	4.97	5.03	5.02	4.98	5.10	4.90	5.39	4.60
SD	1.11	1.10	1.20	1.10	1.18	1.08	1.29	1.08	1.33	.90
<i>F</i> s		5.30*		.14		.09		1.34		24.45***
<b>Promax-Quartimax</b>										
Means	5.05	4.95	5.06	4.94	5.19	4.80	5.20	4.80	5.38	4.62
SD	1.37	1.27	1.28	1.18	1.30	1.11	1.49	1.21	1.41	1.00
<i>F</i> s		.26		.46		5.17*		4.47*		19.32***
<b>Promax-Varimax</b>										
Means	4.91	5.09	5.01	4.99	5.09	4.91	5.11	4.88	5.40	4.59
SD	1.50	1.48	1.57	1.34	1.42	1.26	1.52	1.21	1.43	1.04
<i>F</i> s		.71		.01		.97		1.39		20.81***
<b>Promax-Equamax</b>										
Means	4.90	5.10	5.00	5.00	5.09	4.90	5.11	4.88	5.40	4.60
SD	1.49	1.46	1.55	1.32	1.42	1.26	1.51	1.21	1.43	1.06
<i>F</i> s		.88		.00		1.02		1.44		20.31***

\*  $p < .05$ .

\*\*\*  $p < .001$ .

equamax solutions. The significant *F* ratios indicate that the male *S*s rated themselves as being more ingenious—more skillful and inventive in solving problems—that the female *S*s.

*b. Factor B (Meditativity).* In the quartimax, promax-quartimax, and promax-varimax solutions the means of the female *S*s are higher than those of the male *S*s. In the varimax and equamax solutions the means of the male *S*s are higher than those of the female *S*s, and in the promax-equamax solution the mean of the male *S*s is equal to that of the female *S*s. Since not one of the *F* ratios reached statistical significance at the .05 level, no sex difference was found for factor meditativity.

c. *Factor C (Curiosity)*. In all six factorial procedures the means of the female Ss are higher than those of the male Ss. However, the only statistically significant  $F$  ratio at the .05 level was yielded by the promax-quartimax solution. The significant  $F$  ratio indicates that the female Ss rated themselves as being more curious—more interested in finding out about things, in general—than the male Ss.

d. *Factor D (Alertness)*. In all six factorial procedures the means of the female Ss are higher than those of the male Ss. However, the only statistically significant  $F$  ratio at the .05 level was yielded by the promax-quartimax solution. The significant  $F$  ratio indicates that the female Ss rated themselves as being more alert—more able to size up situations, in general—than the male Ss.

e. *Factor E (Flexibility)*. In all six factorial procedures the means of the female Ss are higher than those of the male Ss, and the  $F$  ratios of all six factorial procedures are statistically significant beyond the .01 level. The significant  $F$  ratios indicate that the female Ss rated themselves as being more flexible—more adjustable to environmental conditions—than the male Ss.

#### D. DISCUSSION AND CONCLUSION

The correlations of the promax-equamax factors, shown in Table 2, range from .28 to .60. They are statistically significant at the .01 level, and indicate that the five factors of the IAL—ingenuity, meditativity, curiosity, alertness, and flexibility—share a certain amount of variance. Factor ingenuity refers to Ss' ability to solve problems skillfully and inventively. Factor meditativity refers to Ss' readiness to become immersed in the inner world of ideas. Factor curiosity refers to Ss' readiness to find out about things. Factor alertness refers to Ss' ability to size up situations, and factor flexibility refers to Ss' adjustment to environmental conditions.

The  $F$  ratios of the five factors of the IAL indicate that the male Ss rated themselves significantly higher than the female Ss on factor ingenuity. However, the female Ss rated themselves significantly higher than the male Ss on factors curiosity, alertness, and flexibility. No significant sex difference was found for factor meditativity.

Pinneau *et al.* (3) have recommended the use of uncorrelated factor scores when performing an analysis of factor variance. In this study, however, both uncorrelated and correlated factor scores were used. Although, neither the orthogonal nor the oblique solutions yielded significant  $F$  ratios on factor meditativity, the oblique solutions as a group yielded significant  $F$  ratios on three of the five factors of the IAL—curiosity, alertness, flexibility—while the or-



thogonal solutions as another group yielded significant  $F$  ratios on only two of the five factors of IAL—ingenuity and flexibility. Moreover, of the six factorial procedures only the promax-quartimax, an oblique solution, yielded significant  $F$  ratios on three of the five factors of the IAL: namely, curiosity, alertness, and flexibility.

The results of Table 3 indicate that the orthogonal and oblique procedures complement each other, since certain  $F$  ratios that did not reach statistical significance in the former procedures reached it in the latter procedures. In general, no factorial procedure was most effective in detecting sex differences among the factors of the IAL. Moreover, whether the statistical significances reported on four of the five factors of the IAL—ingenuity, curiosity, alertness, flexibility—could be attributed to such sources as the  $Ss$ ' social expectations, or their personal preferences, or to the factorial procedures used here, cannot be ascertained from the results of this exploratory study.

Other studies must be carried out, using different samples of male and female  $Ss$ , to investigate (a) the factorial invariance of the IAL; (b) the factorial structure of the IAL under different test instructions; and (c) the influence of different orthogonal and oblique rotations on the estimations of factor scores, and on the statistical significance of the difference between pairs of factor-score means.

#### E. SUMMARY

An iterative factor analysis was performed on the scores obtained from an Intellectual Adjective List (IAL) for a combined sample of undergraduate male and female students ( $n = 202$ ). The initial matrix of significant factors was rotated according to the quartimax, varimax, and equamax procedures. The promax rotation was then performed on each of the three orthogonal solutions obtaining the promax-quartimax, promax-varimax, and promax-equamax solutions. Interpretation of factors was based on the promax-equamax solution and the factors obtained were ingenuity, meditativity, curiosity, alertness, and flexibility.

A simple analysis of factor variance of the factor score means of male and female  $Ss$  on the five factors of IAL, mentioned above, yielded the following results. The male  $Ss$  rated themselves significantly higher than the female  $Ss$  on factor ingenuity; however, the female  $Ss$  rated themselves significantly higher than the male  $Ss$  on factors curiosity, alertness, and flexibility. No significant sex difference was found for factor meditativity. Factor scores were estimated by the multiple regression method and from the quartimax, varimax, equamax, promax-quartimax, promax-varimax, and promax-equamax factorial procedures.

## REFERENCES

1. HENDRICKSON, A. E., & WHITE, P. O. Promax: A quick method for rotation to oblique simple structure. *Brit. J. Stat. Psychol.*, 1964, **17**, 65-70.
2. KAISER, H. F. The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 1958, **23**, 187-200.
3. PINNEAU, S. R., LEVINE, A. J., SCHURR, B. C., & BUTLER, D. C. Analysis of factor variance: One-way classification. *Percept. Motor Skills*, 1966, **23**, 1209-1210.
4. SAUNDERS, D. R. An analytic method for rotation to orthogonal simple structure. Research Bulletin, 53-10, Educational Testing Service, Princeton, N.J., 1953.
5. ———. Trans-varimax: Some properties of the ratiomax and equamax criteria for blind orthogonal rotation. Paper read at the American Psychological Association Convention, St. Louis, Missouri, 1962.
6. SCIORTINO, R. Personality and creative abilities. Doctoral dissertation, Ohio State University, Columbus, Ohio, 1963. No. 64-6958.
7. ———. Factorial analysis of motivational self-ratings from a combined sample of male and female subjects. *Psychol. Rep.*, 1967, **21**, 573-576.
8. ———. Analysis of factor variance of motivational self-ratings by male and female subjects. *J. of Psychol.*, 1968, **69**, 169-174.
9. ———. Factorial study and analysis of factor variance of general adaptability self-ratings by male and female subjects. *J. of Psychol.*, 1969, in press.

*Stevens Institute of Technology*  
*Hoboken, New Jersey 07030*



## FACTORIAL STUDY OF GENERAL ADAPTABILITY SELF-RATINGS BY MALE AND FEMALE SUBJECTS\*

*Stevens Institute of Technology*

---

RIO SCIORTINO

---

### A. INTRODUCTION

A recent factor analysis (5) of the scores of a combined sample of male and female undergraduate Ss on a General Adaptability Adjective List (GAAL) yielded five factors: sociability, energy, attentiveness, quickness, and articulateness. When, in the same study, an analysis of factor variance was performed on the five factors of the GAAL, the results showed that the female Ss rated themselves significantly higher than the male Ss on factors sociability, energy, attentiveness, and quickness; no significant sex difference was found for factor articulateness.

The purpose of this study was to find the factorial structure of the GAAL for a male and a female sample of undergraduate Ss, respectively. Several factorial methods were used and the one that most adequately met the criteria of simple structure was chosen for the interpretation of factors.

### B. METHOD

A male ( $N = 100$ ) and a female sample ( $N = 102$ ) of undergraduate students were administered the GAAL which consists of the 39 items listed in Table 1. The samples, the GAAL, and the test administration procedures were described by Sciortino (5). The findings reported here are based on the two separate groups.

### C. RESULTS

#### 1. Computations

The computations for this study were carried out on the IBM 7094 operated by the Princeton Computer Center and with the P-Stat computer program. For each sample, the iterative factor analysis routine of the P-Stat computer program, described in Sciortino (5), was used to perform a factor analysis on the matrix of item correlations and a matrix of initial significant

---

\* Received in the Editorial Office, Provincetown, Massachusetts, on December 17, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

TABLE 1  
MATRIX OF MEANS AND STANDARD DEVIATIONS FOR SCORES ON  
THE GENERAL ADAPTABILITY ADJECTIVE LIST (GAAL)  
Males ( $N = 100$ ) Females ( $N = 102$ )

Variables	Male sample		Female sample	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
1. Spontaneous	3.14	1.00	3.38	1.02
2. Resilient	3.23	1.04	3.70	.93
3. Expansive	2.99	.82	3.40	.87
4. Cordial	3.30	.77	3.77	.79
5. Amorous	3.36	.97	4.05	.85
6. Warm	3.35	.77	3.98	.83
7. Spirited	3.18	.78	3.63	.83
8. Lively	3.41	.83	3.65	.84
9. Enthusiastic	3.49	.72	3.83	.69
10. Cheerful	3.36	.78	3.80	.92
11. Vigorous	3.38	.84	3.52	.86
12. Quick	3.30	.82	3.33	.72
13. Energetic	3.38	.79	3.50	.85
14. Brisk	3.03	.85	2.87	.97
15. Active	3.61	.80	3.82	.78
16. Wholesome	3.45	.81	3.87	.79
17. Sprightly	3.12	.86	3.31	.93
18. Healthy	3.91	.74	4.15	.93
19. Dynamic	3.08	.76	3.22	.87
20. Agile	3.34	.98	3.40	.97
21. Fluent	2.83	.78	2.95	.97
22. Facile	3.07	.74	3.12	.75
23. Evocative	3.02	.85	3.30	.88
24. Depictive	2.56	.84	2.57	.83
25. Articulate	2.81	.80	3.02	.84
26. Persuasive	3.21	.77	3.64	.98
27. Felicitous	3.07	.90	3.40	.82
28. Eloquent	2.75	.76	2.80	.87
29. Communicative	3.04	.75	3.73	1.04
30. Tactful	3.11	.94	3.15	1.00
31. Responsive	3.18	.82	3.73	.86
32. Leading	3.03	.87	2.87	.98
33. Cooperative	3.46	.80	3.87	.68
34. Attentive	3.30	.72	3.61	.79
35. Tolerant	3.30	.98	3.93	.85
36. Sociable	3.30	.92	3.99	.80
37. Responsible	3.70	.78	3.87	.86
38. Empathetic	3.04	.89	3.48	1.00
39. Congenial	3.46	.81	3.97	.80

factors was obtained. Then, three orthogonal rotations were performed on the matrix of initial significant factors of each sample using the quartimax (3), the varimax (2), and the equamax (4) procedures. For both samples, Table A<sup>1</sup> shows the matrix of item correlations, Table B shows the matrix of initial

<sup>1</sup> The following materials have been deposited with the National Auxiliary Publication Service: Tables A through G contain materials for both the male and female



significant factors, Table C shows the matrix of quartimax factors, Table D shows the matrix of varimax factors, and Table E shows the matrix of equamax factors.

Three promax rotations (1), with  $k = 4$ , were performed on the orthogonal rotations of each sample obtaining the promax-quartimax, the promax-varimax, and the promax-equamax solutions. For both samples, Table F shows the matrix of promax-quartimax factors, and Table G shows the matrix of promax-varimax factors. Table H shows the correlation matrix of the promax-quartimax and promax-varimax factors for the male sample, and Table I shows the correlation matrix of the promax-quartimax and promax-varimax factors for the female sample. For both samples, Table 1 shows the means and standard deviations of the GAAL items and Table 2 shows the matrix of promax-equamax factors. Table 3 shows the correlation matrix of the promax-equamax factors for the male sample, and Table 4 shows the correlation matrix of the promax-equamax factors for the female sample. For both samples, Tables B, C, D, E, F, and G, of the NAPS document, and Table 2, of the present study, show that the total .10 hyperplane count of each of the promax-solutions is higher than that of the corresponding orthogonal solutions. The total .10 hyperplane counts for the promax-quartimax, promax-varimax, and promax-equamax solutions were 137, 143, and 142 respectively in the male sample, and 170, 170, and 167 respectively in the female sample. In this study, as in the previous one (5), the promax-equamax solutions of both samples were used for the interpretation of factors. This could be done because for each sample the difference between the total .10 hyperplane count of the promax-equamax solution and the solution with the highest total .10 hyperplane count is negligible: that is, one unit for the male sample and three units for the female sample.

## *2. Interpretation of Male Sample's Factors*

Interpretation of factors is based on the promax-equamax solution, shown in Table 2, and rests on loadings of .26 or higher ( $p = .01$  if  $N = 100$ ). A

---

samples. Table A, matrix of item correlations ( $39 \times 39$ ) of the GAAL; Table B, matrix of initial significant factors; Table C, matrix of quartimax factors; Table D, matrix of varimax factors; Table E, matrix of equamax factors; Table F, matrix of promax-quartimax factors; Table G, matrix of promax-varimax factors; Table H, correlation matrix of promax-quartimax and promax-varimax factors for the male sample; and Table I, correlation matrix of promax-quartimax and promax-varimax factors for the female sample. Order NAPS Document No. 00281 from ASIS National Auxiliary Publications Service, c/o CCM Information Sciences, Inc., 22 West 34th St., New York, New York, 10001; remitting in advance \$3.00 for photocopies or \$1.00 for microfiche.

TABLE 2  
MATRIX OF PROMAX-EQUAMAX FACTORS FOR SCORES ON THE GENERAL ADAPTABILITY ADJECTIVE LIST (GAAL)  
Males ( $N = 100$ ) Females ( $N = 102$ )

Variables	Promax-Equamax factors							
	Male sample				Female sample			
	A	B	C	D	E	F	G	H
1. Spontaneous	-16	00	26	23	10	21	-13	03
2. Resilient	-04	27	-08	44	-10	-10	01	-04
3. Expansive	09	-01	05	42	07	-15	09	-21
4. Cordial	00	60	-22	10	-05	00	20	09
5. Amorous	-01	00	02	57	02	-09	-07	-01
6. Warm	08	14	08	58	-11	-13	-07	04
7. Spirited	57	-09	06	06	-06	22	-07	-14
8. Lively	59	04	15	-02	-03	09	-07	07
9. Enthusiastic	05	00	29	-04	-03	14	27	06
10. Cheerful	09	40	24	-05	-06	05	20	14
11. Vigorous	-01	-06	59	12	06	-07	-02	14
12. Quick	33	03	-04	39	-12	17	-12	48
13. Energetic	28	-11	56	-01	17	-15	04	15
14. Brisk	47	03	10	16	01	-04	-14	58
15. Active	02	00	58	00	08	-14	-01	12
16. Wholesome	34	-14	12	12	00	-03	35	-06
17. Sprightly	53	00	-02	03	15	09	05	08
18. Healthy	54	-03	02	-01	-02	04	18	-16
19. Dynamic	14	-05	40	-11	-11	15	38	03
20. Agile	30	23	08	02	08	01	00	09
21. Fluent	-02	26	10	-09	53	04	-21	15
22. Facile	12	15	24	15	-02	18	-01	54
23. Evocative	-13	-01	03	29	23	18	11	-08
24. Depictive	-02	-08	22	00	67	-13	-08	12
25. Articulate	09	-01	03	-12	60	02	00	24
26. Persuasive	10	08	-09	-09	04	56	00	00
27. Felicitous	-06	41	17	16	13	-03	05	09
28. Eloquent	-06	07	-05	-11	28	3+	14	17

TABLE 2 (continued)

Variables	Promax-Equamax factors														
	Male sample							Female sample							
	A	B	C	D	E	F	G	A	B	C	D	E	F	G	H
29. Communicative	04	57	-10	-13	23	21	-18	-05	-05	52	01	07	11	06	06
30. Tactful	04	-11	-26	15	37	21	01	00	11	05	-10	-02	10	51	03
31. Responsive	12	00	-16	27	06	13	17	-08	06	09	54	17	-13	-04	-01
32. Leading	18	06	-12	-15	-07	65	06	22	-13	10	02	-01	14	19	19
33. Cooperative	08	29	08	-21	-01	-06	42	-08	64	-06	-01	02	05	10	00
34. Attentive	10	11	-12	-02	29	-31	37	-03	53	-17	23	-03	01	14	-04
35. Tolerant	-15	24	-06	02	-16	02	56	-13	39	02	19	13	00	-03	07
36. Sociable	-04	46	05	10	-08	09	-06	-04	05	43	20	07	-14	24	02
37. Responsible	-03	-17	06	-10	-09	16	50	02	48	05	-05	-07	-22	07	14
38. Empathetic	-27	-04	03	18	19	07	36	05	05	-17	49	11	24	05	-17
39. Congenial	-08	61	00	21	-11	-03	15	-18	38	46	28	-18	-07	21	-08
.10 h.c.	22	21	22	17	22	19	19	20	21	21	21	22	18	20	24

Note: Initials h.c. stand for hyperplane count. Decimal points omitted.

TABLE 3  
CORRELATION MATRIX OF PROMAX-EQUAMAX FACTORS FOR SCORES ON  
THE GENERAL ADAPTABILITY ADJECTIVE LIST (GAAL)  
Males ( $N = 100$ )

Factors	Promax-Equamax factors					
	A	B	C	D	E	F
A (Liveliness)						
B (Congeniality)	.41*					
C (Vigor)	.44*	.37*				
D (Warmth)	.21	.29*	.26*			
E (Depictiveness)	.08	.11	-.02	.35*		
F (Leadership)	.17	.18	.30*	.45*	.42*	
G (Tolerance)	.27*	.06	.03	.26*	.26*	.22

Note: Decimal points omitted.

\*  $p = .01$ .

TABLE 4  
CORRELATION MATRIX OF PROMAX-EQUAMAX FACTORS FOR SCORES ON  
THE GENERAL ADAPTABILITY ADJECTIVE LIST (GAAL)  
Females ( $N = 102$ )

Factors	Promax-Equamax factors						
	A	B	C	D	E	F	G
A (Sprightliness)							
B (Cooperativeness)	.45*						
C (Communicativity)	.52*	.37*					
D (Responsiveness)	.40*	.32*	.42*				
E (Amorosity)	.22	.39*	.31*	.26*			
F (Fluency)	.26*	.14	.20	.23	.11		
G (Promptness)	-.02	.17	-.04	.02	.15	-.09	
H (Briskness)	.46*	.18	.29*	.37*	.04	.21	-.19

Note: Decimal points omitted.

\*  $p = .01$ .

factor, whenever possible, was assigned the name of the variable that has the highest loading with it.

Factor A (Liveliness) has loadings arranged as follows: lively, .59; spirited, .57; healthy, .54; sprightly, .53; brisk, .47; wholesome, .34; quick, .33; agile, .30; energetic, .28; and empathetic, -.27. This factor refers to Ss' liveliness or vivacity. However, the negative loading of item empathetic suggests that Ss' liveliness detracts from their ability to empathize with others.

Factor B (Congeniality) has loadings arranged as follows: congenial, .61; cordial, .60; communicative, .57; sociable, .46; felicitous, .41; cheerful, .40; cooperative, .29; resilient, .27; and fluent, .26. This factor refers to Ss' congeniality or disposition to be friendly and sympathetic with others.

Factor C (Vigor) has loadings arranged as follows: vigorous, .59; active,

.58; energetic, .56; dynamic, .40; enthusiastic, .29; spontaneous, .26; and tactful, —.26. This factor refers to Ss' vigor or active strength. However, the negative loading of item tactful suggests that Ss' vigor detracts from their success in dealing with others.

Factor D (Warmth) has loadings arranged as follows: warm, .58; amorous, .57; resilient, .44; expansive, .42; quick, .39; evocative, .29; and responsive, .27. This factor refers to Ss' warmth or affectionate concern for others.

Factor E (Depictiveness) has loadings arranged as follows: depictive, .67; articulate, .60; fluent, .53; tactful, .37; and attentive, .29. This factor refers to Ss' depictiveness or facility in describing events.

Factor F (Leadership) has loadings arranged as follows: leading, .65; persuasive, .56; eloquent, .34; and attentive, —.31. This factor refers to Ss' leadership, or ability to direct people's activities. However, the negative loading of item attentive suggests that Ss' leadership detracts from their consideration for others.

Factor G (Tolerance) has loadings arranged as follows: tolerant, .56; responsible, .50; cooperative, .42; dynamic, .38; attentive, .37; empathetic, .36; wholesome, .35; and enthusiastic, .27. This factor refers to Ss' tolerance or liberal attitude toward others.

### 3. *Interpretation of Female Sample's Factors*

Interpretation of factors is based on the promax-equamax solution, shown in Table 2, and rests on loadings of .25 or higher ( $p = .01$  if  $N = 102$ ). A factor, whenever possible, was assigned the name of the variable that has the highest loading with it.

Factor A (Sprightliness) has loadings arranged as follows: sprightly, .68; spirited, .64; lively, .61; energetic, .46; agile, .42; dynamic, .37; healthy, .33; active, .31; resilient, .29; vigorous, .27; and enthusiastic, .25. This factor refers to Ss' sprightliness or tendency to move rapidly and lightly.

Factor B (Cooperativeness) has loadings arranged as follows: cooperative, .64; attentive, .53; responsible, .48; wholesome, .44; tolerant, .39; congenial, .38; cheerful, .37; cordial, .29; healthy, .29; and enthusiastic, .28. This factor refers to Ss' cooperativeness or readiness to act jointly with others.

Factor C (Communicativity) has loadings arranged as follows: communicative, .52; congenial, .46; spontaneous, .44; sociable, .43; persuasive, .35; expansive, .29; and eloquent, .26. This factor refers to Ss' communicativity or readiness to talk with others.

Factor D (Responsiveness) has loadings arranged as follows: responsive,



.54; empathetic, .49; evocative, .41; expansive, .37; brisk, —.30; congenial, .28; and eloquent, .25. This factor refers to Ss' responsiveness or readiness to react to others in sympathy. However, the negative loading of item brisk suggests that Ss' responsiveness has a certain hampering effect on their quickness of movement.

Factor E (Amorosity) has loadings arranged as follows: amorous, .70; warm, .59; and persuasive, .27. This factor refers to Ss' amorosity or readiness to love others.

Factor F (Fluency) has loadings arranged as follows: fluent, .73; articulate, .56; depictive, .46; healthy, .27; and evocative, .26. This factor refers to Ss' fluency or ability to express themselves easily.

Factor G (Promptness) has loadings arranged as follows: quick, .56; tactful, .51; cordial, .46; and resilient, .38. This factor refers to Ss' promptness or tendency to respond rapidly, though tactfully, to others.

Factor H (Briskness) has loadings arranged as follows: brisk, .58; facile, .54; and quick, .48. This factor refers to Ss' briskness or spryness in movement.

#### D. DISCUSSION AND CONCLUSION

The correlations of the male Ss' seven factors—liveliness, congeniality, vigor, warmth, depictiveness, leadership, tolerance—shown in Table 3, range from —.02 to .45, and those of the female Ss' eight factors—sprightliness, cooperativeness, communicativity, responsiveness, amorosity, fluency, promptness, briskness—shown in Table 4, range from —.09 to .52. In each sample, with the exception of the female Ss' factor promptness, all factors correlate significantly with certain other factors and share, therefore, a certain amount of variance.

For the male sample, the factors indicate that the Ss' interaction with others is generally marked by sympathy, concern, liberality, as well as a certain facility of communication. However, they also indicate that Ss' vivacity, vigor, and readiness to lead others might, on occasion, interfere with their success in interpersonal relations.

For the female sample, the factors indicate that the Ss' interaction with others is generally marked by cooperativeness, talkativeness, sympathy, love, as well as a certain ease of expression. However, they refer also to Ss' briskness of movement which, on occasion, might be attenuated, or restrained, by their sympathetic responses to others.

The results reported here indicate that, under the conditions of this investigation, the factorial structures of the GAAL items, for the male and female samples, reflect sex differences. The influence of sex has been observed in the

number of factors extracted—seven for the male and eight for the female sample—as well as in the item composition and item loadings of the respective factors of the two samples. By contrast, the five factors—sociability, energy, attentiveness, quickness, and articulateness—extracted from the scores of the combined sample of male and female Ss on the GAAL items (5) have more generality than the seven factors of the male sample and the eight factors of the female sample reported here.

These three analyses suggest that, if one is interested in obtaining factors of a general nature, then a factorial study in which both male and female samples are combined is indicated. However, if one is interested in obtaining factors that better describe the general adaptability configuration of either sex, a separate factor analysis for either sex is in order.

### E. SUMMARY

Two separate iterative factor analyses, performed on the scores of a General Adaptability Adjective List (GAAL) by a male sample ( $N = 100$ ) and a female sample ( $N = 102$ ), yielded two matrices of initial significant factors. For each sample, the initial matrix of significant factors was rotated according to the quartimax, varimax, and equamax procedures. Then the promax rotation was performed on the orthogonal solutions of each sample obtaining the promax-quartimax, the promax-varimax, and the promax-equamax solutions. Interpretation of factors was based on the promax-equamax solutions of each sample. The factors obtained with the male sample were, liveliness, congeniality, vigor, warmth, depictiveness, leadership, and tolerance. The factors obtained with the female sample were, sprightliness, cooperativeness, communicativity, responsiveness, amorosity, fluency, promptness, and briskness.

### REFERENCES

1. HENDRICKSON, A. E., & WHITE, P. O. Promax: A quick method for rotation to oblique simple structure. *Brit. J. Stat. Psychol.*, 1964, **17**, 65-70.
2. KAISER, H. F. The variance criterion for analytic rotation in factor analysis. *Psychometrika*, 1958, **23**, 187-200.
3. SAUNDERS, D. R. An analytic method for rotation to orthogonal simple structure. Research Bulletin, 53-10, Educational Testing Service, Princeton, New Jersey, 1953.
4. ———. Trans-varimax: Some properties of the ratiomax and equamax criteria for blind orthogonal rotation. Paper read at the American Psychological Association Convention, St. Louis, Missouri, 1962.
5. SCIORTINO, R. Factorial study and analysis of factor variance of general adaptability self-ratings from a combined sample of male and female subjects. *J. of Psychol.*, in press.

*Stevens Institute of Technology*  
*Hoboken, New Jersey 07030*



## BRIGHTNESS MATCHING UNDER INTERMITTENT STIMULATION\*<sup>1</sup>

*Department of Psychology, Michigan State University*

JONATHAN D. KAZSUK AND S. HOWARD BARTLEY

### A. INTRODUCTION

#### 1. *Purpose*

The purposes of this paper are: (a) to point out that as the rate of intermittent photic stimulation is reduced from CFF, several rate ranges produce distinctly different sensory phenomena; (b) to report findings that show that all segments of the overall range cannot be responded to in the same manner; (c) to point out that in certain cases in the literature the label of brightness enhancement has been misapplied, because the technical definition has been ignored; and (d) to show why these facts nullify certain criticisms in the literature regarding brightness enhancement.

Point *a* simply reminds many readers of what they know already, whereas the remaining points have to do with clearing up some oversights and naiveté, and presenting some additional evidence.

#### 2. *Sensory End Results*

At least four types of sensory phenomena occur as rate of intermittency is extensively manipulated. The exact frequencies at which they occur are affected by pulse length, intensity, target size, wavelength, image position, and photic surround conditions of the target. Schneider and Bartley (19) have described four basic categories as follows: (a) One in which distinct light and dark target phases alternate, their relative prominence depending upon the relative duration of the photic pulses and the null periods between. (b) A range in which the two phases are not separate but rather create a greatly fluctuating brightness sequence (course flicker). (c) A range of minor fluctuations (fine flicker) superimposed upon a steady brightness, a range

\* Received in the Editorial Office, Provincetown, Massachusetts, on January 9, 1969, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This investigation was supported by Public Health Service Fellowship 5 F1 MH-34,507 from the National Institute of Mental Health and by Public Health Service Research Grant NB 05260 from the National Institute of Neurological Diseases and Blindness.

within which the observer is fairly able to compare the brightness of the field with the brightness of an adjacent steady one. (d) A range throughout which the fluctuations are absent and the field possesses a steady brightness. This is often called fusion.

Ranney (17) found that the range between fusion and single flashes could be divided into course flicker, fine flicker, ripple, and flutter. The photic pulse rates for the boundaries between these depended upon pulse-to-cycle fraction (PCF) and to some extent upon target luminosity. The boundary between course and fine flicker was in the neighborhood of five pulses per second.

The major question that arises in matching the sensory end results of intermittent stimuli with those of steady ones is whether it is possible to match the brightness results produced at all rates, despite the distinctions just pointed out. First of all, is there a brightness at all rates that can be used in the matching procedure, and second, can the results from categories (ranges) *a*, *b*, and *c* be lumped together as though the sensory end results were all phenomenologically similar? Obviously, with category *a*, the flashes stand out as distinct and might be expected to be matched with the brightness of the steady standard target. But, with flashes distinct from the dark periods, do the null periods play a role in the brightness of the flashes? At least it can be said that the two phases cannot be phenomenologically averaged by the observer.

Within category *b*, both phases of the stimulus cycle are occupied by an illuminated field of some brightness. That is, bright phases and less bright phases alternate. Again, no averaging is quite possible by the observer. He may attempt to match the brightness of the bright phases with the standard target, or he may attempt to do this for the less bright phases.

With category *c*, averaging is possible. In fact, the fluctuation is superimposed upon a salient steady component. The fluctuations are trivial in comparison with this steady component and are not separable for use as in category *b*. It was this category (category *c*) that was used by Bartley (3) in defining brightness enhancement in the first place.

It should be axiomatic that when one uses a range of stimulus conditions which produce a sequence of diverse sensory end results in different segments of the range, all of the end results can be grouped in the same category only by reason of some common denominator among them. The feature of the phenomena in category *c* (one segment of the range) which we have already defined, is that the field possesses a steady aspect that can be compared to the steady field produced by steady stimulation (standard target). As one lowers



the rate of stimulus intermittency to where only one component in the brightness fluctuation is usable (categories *b* and *c*), then an essential phenomenological boundary has been passed. If one ignores this boundary, it is like grouping horses and cows together and calling them all cows.

In the literature, we find that categories *a* and *b* (not appropriate for testing brightness enhancement) have been utilized to test brightness enhancement. While these studies tested brightness, they did not test brightness enhancement. The term brightness enhancement arose in the first place by definition. Usage should either follow the definition or should explicitly show need for revision or extension. This matter has already been discussed by Schneider and Bartley (19).

### 3. *The Theory Supposedly Tested*

The context which has given the most relevance and significance to the facts of brightness enhancement is the alternation of response theory (5, 6, 7). This theory is not simply a statement expressing an isolated hypothesis regarding a single phenomenon, such as brightness enhancement, but instead, it is an extensive picture of certain temporal features of activity in the central end of the optic pathway. This theory accounts for the unexpected brightness which otherwise would be left unexplained by traditional concepts, or be a genuine paradox, a case of "getting something (brightness) for nothing."

It has been frequently implied, if not stated, in the literature that the theory requires that maximal brightness enhancement be produced by an intermittency rate of 10 cps. No statement of the theory was ever made which required maximum brightness enhancement to be at approximately 10 cycles per second under all conditions. A fuller understanding of the theory to be gained by reading the many papers written on its various expressions would not lead to such an expectation. Those workers who have found brightness maxima at other rates have not succeeded in testing the theory and gaining evidence against it, as they thought they had.

Under the conditions initially used and in certain other investigations since, the stimulus intermittency rate which produced maximum brightness enhancement was in the neighborhood of 10 cycles per second (2, 3, 8, 9, 14), while in other investigations it was about five per second. In the literature in general, rates for maximum have ranged from one to 12 cycles per second or even higher (1, 4, 9, 10, 11, 12, 13, 15, 16, 18). But since not all of these conditions were actually those for studying brightness enhancement, they do not prove or disprove what rates produce *maximum* enhancement.

For example, rates of intermittency low enough to produce the sensory end results of our category *a* would, according to the alternation of response theory, produce a brightness which is greater than higher rates of intermittency producing a sensory end result characteristic of category *c*. The confusion follows from not fully recognizing that a brightness match of targets in category *a* is a match of flashes of light with the steady comparison target, whereas intermittent photic stimulation for category *c* produces an appreciable steady brightness component appropriate for matching with the steady target.

Up until the study of Colgan (12), no researcher reported giving specific instruction to the observers making intermittent brightness matches. Hence the problem just stated was never explicitly faced. Colgan showed that if an observer is instructed to match to the "on" phase of the brightness fluctuations, there will be a significant increase in the obtained brightness levels. This was a confirmation of the idea that one can choose one component or the other when the category *b* condition is used.

Schneider and Bartley (19) compared subjective brightness matching criteria over a range of frequencies from 1 to 20 cps, with a 20 msec. pulse. They found that the standard deviations were greater for 3, 4, 5, and 6 cps than for the higher rates used. They suggested that the greater variance was produced by the observer's shifting his brightness match "criterion" condition from one session to another at these rates. It is easy to see that the observer is confronted with several possibilities as to "what" brightness to be dealing with, hence it is not surprising that the observers tended to shift from time to time.

## B. METHOD

The apparatus was essentially the same as described in Schneider and Bartley (19) and need not be described again here. The difference pertained to target arrangement. The targets, one seen above the other, subtended a visual angle of  $43'$  of arc when viewed from a distance of 80 cm. The interspace between them was  $2^{\circ}9'$ .

Two male graduate students, including one of the present authors, served as the observers. Each had many hours of experience serving in subfusional intermittent brightness matching experiments.

The intensity of the intermittent target was held constant at 1.8 cu. ft.<sup>2</sup>, while the intensity of the steady target could be manipulated from 3.8 cu. ft.<sup>2</sup> to .1 cu. ft.<sup>2</sup> by the movement of the neutral density wedge. The current for both the steady and intermittent glow tubes was monitored and held constant at 25 ma. and made the appearance of the glow tubes nearly achromatic.

Before collecting data, each observer practiced the instructions for several weeks in order to become familiar with the tasks.

For all conditions, the observer, using monocular vision, matched the brightness of the steady target to the brightness of one of the three aspects of the brightness fluctuation of the intermittent target.

The *on-match* instructional set meant that the observer attended and matched to the bright phase of the fluctuation. The *mean-match* instructional set meant that the observer attended and matched to a derivative of the brightness fluctuation. The *off-match* instructional set required that the observer attend and match to the dark phase of the fluctuation.

A total of 3060 observations were recorded for each observer. In connection with the three instructional sets, three pulse lengths were used. For the pulse length of 20 ms., frequencies of 45, 25, 20, 15, 12, 10, 8, 6, 4, 2, 1, .75, and .50 cps were used. For the pulse length of 40 ms., the frequencies of 45 and 25 cps were not used. For the pulse length of 60 ms., the frequency of 20 cps was also not used. Ten observations (five ascending and five descending) were made in each of three sessions for a total of 30 observations for all conditions. In addition to the 3060 intermittent observations, each observer made a total of 540 brightness matches with both targets at steady luminance. Ten steady observations were recorded at the beginning of each session and an additional 10 were recorded at the end of each session.

One pulse length with the various rates presented in random order was used over three sessions for the three instructional sets before the next pulse length was used.

Ten minutes for adapting to the target were given before each session.

All values obtained with the neutral density wedge were converted into the percentage of transmission (PT) required for a match. The mean of 30 observations made under a particular condition was used as the PT required to make brightnesses of the two targets appear equal. The mean of the 540 steady matches was calculated as the steady match.

### C. RESULTS

The direct brightness results are given in Figure 1 in which the responses of two observers for three pulse durations and a range of stimulus intermitencies under three instructional conditions are shown. Brightness is expressed in the form of an index, since it was indicated only in terms of making the intermittent target match the steady one. The index is expressed in per cent. The optical wedge setting for matching the two targets when the luminance in both is steady is taken as 100 per cent.

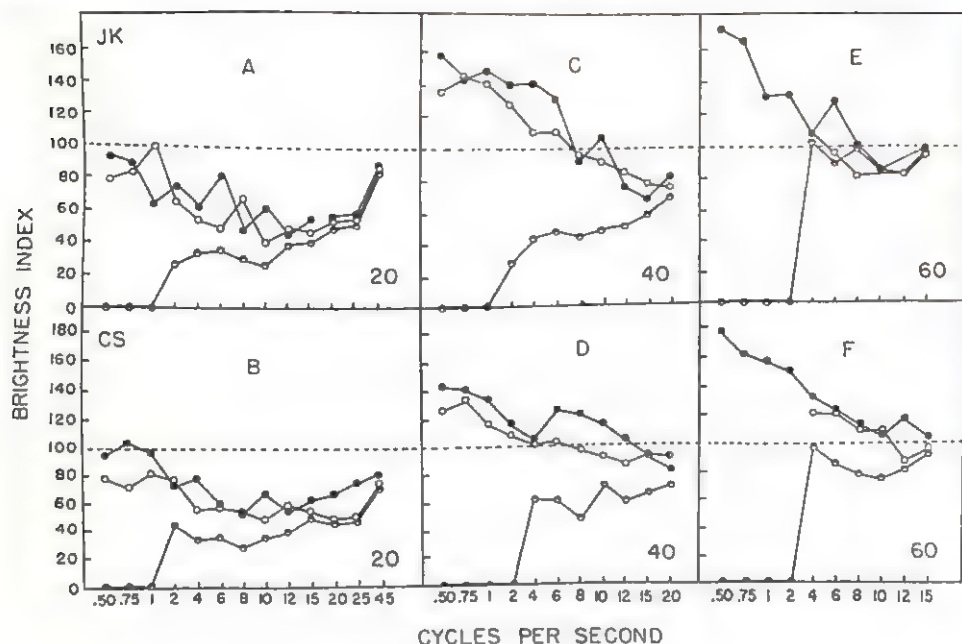


FIGURE 1

RELATION OF BRIGHTNESS TO THE RATE OF INTERMITTENCY OF A PHOTIC STIMULUS

The upper row (ACE) is for observer JK; the lower row (BDF), for observer CS. The numbers 20, 40, and 60 represent the durations of the photic pulses in milliseconds. The dotted lines represent the brightness index for steady stimulation.

Two sorts of results are to be taken into account: first, the protocols of the observers in following the instructions; and second, the actual brightness indexes provided by their responses.

It will be evident from A, B, C, and D of the figure that responses were given for the "mean" instructions even for the very lowest rates of intermittency, except when 60 msec. pulses were used. On the surface, this would seem that it is appropriate to use very low rates to obtain data regarding brightness enhancement, with the use of Bartley's original definition of the term. If one takes the mean to indicate an arithmetic mean of the values obtained by averaging the responses obtained under the "on" and "off" conditions, it is obvious that this is not what was obtained. The so-called mean results follow fairly closely the values for the "on" instructions. The observers report that they were literally unable to see a mean brightness when the rates became low, but were able to make a response which they believed at the time was a mean between the light and dark phases of the fluctuation. It must be concluded that this "mean" was not a true direct sensory product



but a kind of judgment, and thus one that cannot be used for the same purposes as a true mean in the phenomenological sense. By the time that pulse duration was lengthened to 60 msec., the observers were no longer able to respond under "mean" instructions when the intermittency rate declined below 4 cps. This was the same rate at which the "off" readings last demonstrated any brightness at all. For a pulse duration of 40 msec., brightness became zero for one of the observers at a rate of 2 cps. For the other, the last readings were obtained at 1 cps.

The data for the higher rates of intermittency suffer from the opposite difficulty; namely, that the mean response is probably becoming the easiest of the three to make. In fact, for the shortest pulse duration (20 msec.), the "on," "off," and "mean" readings are essentially alike at the highest rates used. And, for the longest pulse durations, they are only slightly different. Furthermore, there is considerable fluctuation in the relative levels for each of the instructional sets. This attests to some difficulty in making the readings.

At rates below 4 cps for the 60 msec. pulse, and below 1 cps for the 20 msec. pulse, and below 2 for JK and below 4 for CS for the 40 msec. pulse, no brightness was reported for the "off" instructional set.

For the "on" instructional set, the brightness readings ascended from below the level of brightness for the steady matches, to those having an index of 160 to 180 per cent above for the 40 and 60 msec. pulses as rate declined. Thus it is easy to see why some reports of "brightness enhancement" increasing steadily, down to rates of 1 or 2 cps, have appeared in the literature. They would have been produced by observers trying to respond to the "mean" or to the "on" (bright) phase of the intermittency cycle.

The observers in this investigation found no difficulty in responding under the "on" and "off" instructions when rates were low. When a pulse duration of 60 msec. was used, it became apparent that the observers were totally unable to make a "mean" match of any sort, and also unable to see any brightness under "off" instructions, when rate declined to 2 cps.

In the present study when brightness of the intermittent target rose above the brightness for the steady one, it was at rates definitely below 10 cps and brightness rose as rate declined. This was despite the fact that the authors used luminosity levels below those used to obtain brightness enhancement (at the proper intermittency rates).

Bartley (4) pointed out that when intermittent targets of low luminosities do produce brightnesses above the level of steady targets, this rise in brightness is to be expected at low rates of intermittency. That brightnesses greater



than those produced by the steady target were obtained (at the low rates of intermittency), is a confirmation of Bartley's position instead of being evidence against the interpretation he earlier made regarding brightness maxima.

#### D. SUMMARY AND CONCLUSIONS

The present study was made to deal with certain confusions regarding brightness obtained with low rates of intermittent photic stimulation. In so doing, it was shown that, as rates become low, a point is reached at which the intermittent target is no longer seen as one having a steady brightness component to match with the steady standard target. It becomes a field of definitely fluctuating brightness. The observer's problem then becomes very different than before. He now is faced with choosing whether to try to make a "mean" or average reading of the fluctuation or to deal with the low brightness phases or the high brightness phases of the fluctuation. In any case, he is no longer doing what he was able to do up to this point.

The present findings pertain to what happens under the three possible sets of instructions—high brightness ("on"), low brightness ("off"), or an attempted "mean" set of readings.

The data show that under some conditions, "mean" readings are impossible (at low rates), and also that "off" readings are also impossible (at high rates). The results replicate, in principle, a variety of findings in the literature and at the same time show how these were misinterpreted. The data that were added, further the demonstration that findings from all conditions (rates, etc.) cannot be compared as if they pertained to a single set of sensory phenomena.

#### REFERENCES

1. ALPERN, M., & SUGIYAMA, S. Photic driving of the critical flicker frequency. *J. Opt. Soc. Amer.*, 1961, **51**, 1379-1385.
2. BALL, R. J. An investigation of chromatic brightness enhancement tendencies. *Amer. J. Optom. & Arch. Acad. Optom.*, 1964, **41**, 333-361.
3. BARTLEY, S. H. Subjective brightness in relation to flash rate and the light-dark ratio. *J. Exper. Psychol.*, 1938, **23**, 313-319.
4. ———. Intermittent photic stimulation at marginal intensity levels. *J. of Psychol.*, 1951, **32**, 217-223.
5. ———. Some facts and concepts regarding the neurophysiology of the optic pathway. *Arch. Ophthal.*, 1958, **60**, 775-791.
6. ———. A clarification of some of the procedures and concepts involved in dealing with the optic pathway. In R. Jung & H. Kornhuber (Eds.), *The Visual System: Neurophysiology and Psychophysics*. Heidelberg, Germany: Springer-Verlag, 1961.
7. BARTLEY, S. H., & NELSON, T. M. Some relations between sensory end results and neural activity in the optic pathway. *J. of Psychol.*, 1963, **55**, 121-143.

8. BARTLEY, S. H., PACZEWITZ, G., & VALSI, E. Brightness enhancement and the stimulus cycle. *J. of Psychol.*, 1957, **43**, 187-193.
9. BARTLEY, S. H., NELSON, T. M., & SOULES, E. M. Visual acuity under conditions of intermittent illumination productive of paradoxical brightness. *J. of Psychol.*, 1963, **55**, 153-163.
10. BATTIG, W., VOSS, J., & BROGDEN, W. The effect of frequency of intermittence upon perceived brightness. *J. Comp. & Physiol. Psychol.*, 1956, **50**, 61-63.
11. BOURASSA, C. M., & BARTLEY, S. H. Some observations of the manipulation of visual acuity by varying the rate of intermittent stimulation. *J. of Psychol.*, 1965, **59**, 319-328.
12. COLGAN, C. M. The effect of observational technique on brightness enhancement. *Amer. J. Psychol.*, 1965, **78**, 471-475.
13. GULICK, W. L. Brightness enhancement with microsecond pulses. *Psychol. Rec.*, 1961, **11**, 249-256.
14. HALSTEAD, W. C. A note on the Bartley effect in the estimation of equivalent brightness. *J. Exper. Psychol.*, 1941, **28**, 524-528.
15. LINDSLEY, D. M. Some facts and concepts regarding the neurophysiology of the optic pathway. (In discussion of article by S. H. Bartley) *Arch. Ophthalm.*, 1958, **60**, 775-791.
16. NELSON, T. M., BARTLEY, S. H., & FORD, Z. The brightness of part-spectrum targets stimulating the eye parafoveally at intermittency rates below fusion. *J. of Psychol.*, 1963, **55**, 387-396.
17. RANNEY, J. E. Perceptions of subfusional photic intermittencies under various intensities and pulse-to-cycle fractions. Unpublished Doctoral dissertation, Michigan State University, East Lansing, Michigan, 1964.
18. REIDMEISTER, C., & GRUSSER, O. J. Flimmerlichtuntersuchungen an der Katzenretina. *I. Z. Biol.*, 1959, **111**, 241-253.
19. SCHNEIDER, C. W., & BARTLEY, S. H. Changes in sensory phenomena and observer criteria at low rates of intermittent photic stimulation. *J. of Psychol.*, 1966, **63**, 53-66.

*Department of Psychology*  
*Michigan State University*  
*East Lansing, Michigan 48823*









